Insurance-Induced Moral Hazard: A Dynamic Model of Within-Year Medical Care Decision Making Under Uncertainty

Christopher J. Cronin[1]

University of Notre Dame

February 8, 2018

Abstract

This study quantifies the moral hazard effect of health insurance on medical expenditure by estimating a dynamic model of within-year medical care consumption that allows for insurance selection, endogenous health transitions, and individual uncertainty about medical care prices in an environment where insurance has non-linear cost-sharing features. The results suggest that moral hazard accounts for 53.1 percent, on average, of total annual medical expenditure when insured. This moral hazard effect is found to be significantly different, and generally larger, than that produced by an alternative model that is representative of the annual medical care decision-making models found in the literature.

JEL Classification: C61, D81, G22, I12, I13

Keywords: Dynamic Discrete Choice, Choice Under Uncertainty, Risk Aversion, Health Insurance, Within-Year Decision Making, Health Production

---

[1]I would like to thank the coeditor and two anonymous referees; my dissertation advisor Donna Gilleskie; as well as Chuck Cortemanche, Michael Darden, David Guilkey, Randy Ellis, Michael Grossman, Matt Harris, Clement Joubert, Brian McManus, Tiago Pires, Dan Rees, Forrest Spence, Steve Stern, and Helen Tauchen for their helpful comments. I also benefited from conversations with seminar participants at numerous conferences and universities. Email: ccronin1@nd.edu. This research was conducted at the Triangle and Chicago Census Research Data Centers, and support from the Agency for Healthcare Research and Quality (AHRQ) is acknowledged. The results and conclusions in this paper are those of the author and do not indicate concurrence by AHRQ or the Department of Health and Human Services.

1 Introduction

Economic theory suggests that health insurance may increase medical care consumption above the socially optimal level (Arrow, 1963; Pauly, 1968). The incentives that elicit this increase in consumption are often referred to as moral hazard (Cutler and Zeckhauser, 2000).[1] Empirical studies tend to estimate moral hazard effects using models that aggregate medical care decisions up to the annual level (Cardon and Hendel, 2001; Khwaja, 2010; Einav et al., 2013; Bajari et al., 2014; Kowalski, 2015). In this paper, I study insurance-induced moral hazard using a dynamic, stochastic model of within-year medical care decisions. The within-year decision-making (WYDM) model more accurately captures the data generating process by relaxing several assumptions made frequently in the literature. Specifically, the model allows for endogenous health transitions, variation in medical care prices, and individual uncertainty within a health insurance year.[2] The primary objective of this research is to determine whether the WYDM model produces different, and more accurate, moral hazard estimates than an annual decision-making model.

Allowing an individual's medical care optimization problem to evolve endogenously over the course of a health insurance year has implications for the analysis of moral hazard. For example, if medical care consumption affects future health transitions, then a generous insurance plan, relative to a less generous plan, may lead an individual to consume more medical care early in the year, experience health improvement, and spend less late in the year. Additionally, if an individual's medical care decisions are determined in part by his forecast of future health and medical expenditure, then health insurance can affect spending patterns through its impact on the forecast horizon, in addition to its impact on out-of-pocket prices. The cumulative effect that these mechanisms have on moral hazard is ex-ante ambiguous, which motivates this empirical investigation.

The principle innovation of this research is the development and estimation of a WYDM model, which is mo-

[1] The term *moral hazard* is used rather loosely in the health economics literature. In this paper, I refer to moral hazard as a change in the incentives for purchasing medical care due to insurance acquisition. I focus on several *effects* of moral hazard, such as the change in total annual medical expenditure that results from insurance acquisition.

[2] The *price* of medical care can refer to any of three different values for insured individuals in the US. A *list* price can be thought of as the theoretical market price for care. A *transaction* price is the sum of the insurer's and insured's payments and is typically lower than a list price as insurance companies negotiate with medical care providers for reduced rates. An *out-of-pocket* price is the share of a transaction price paid by an insured individual. In this manuscript, any reference to *price* refers to a transaction price. List prices are rarely relevant to an individual's decision making, which is the focus of this paper. When discussed, out-of-pocket prices are referenced accordingly.

tivated by theoretical (Grossman, 1972; Keeler et al., 1977) and empirical (Gilleskie, 1998; Khwaja, 2010; Blau and Gilleskie, 2008) models of health production and medical care demand. An individual's optimization problem consists of an annual health insurance decision, followed by a sequence of monthly medical care decisions made over the course of a health insurance year. I model monthly medical care decisions to allow the unique benefits and costs associated with the timing of unexpected illness and dynamic insurance cost-sharing features to impact behavior within the model. Within each month, an individual responds to an endogenous, stochastic health event by consuming medical care. The anticipated primary benefit of medical care consumption is improved future health. The anticipated primary cost is financial (i.e., a decrease in the current consumption of non-medical goods). When health insurance has dynamic cost-sharing features (i.e., a deductible and maximum out-of-pocket expenditure (MOX) level), an additional benefit of medical care consumption is lower future out-of-pocket prices.[3] Importantly, the model assumes that individuals are forward-looking, meaning they consider the impact that their actions have on the probability of future outcomes (e.g., health and prices) when making decisions. Evidence of such forward-looking health behavior can be found in Arcidiacono et al. (2007).

The WYDM model is expressed as a dynamic programming problem (DPP), which is solved via backwards recursion and estimated via maximum likelihood (MLE) using employer-employee matched data from the Medical Expenditure Panel Survey (MEPS). The use of nationally representative survey data is unique in this literature, as most recent research has analyzed firm-level claims data. The MEPS data include dynamic medical care and health information for individuals working for different firms that offer a wide variety of observable health insurance alternative sets. Variation in insurance alternative sets aids in the identification of moral hazard effects as it ensures that observationally equivalent individuals hold different plans and, thus, face a different sequence of out-of-pocket medical care prices over the insurance year. Another advantage of the MEPS data is that illness episodes are reported even when medical care is not consumed, which allows endogenous illness transitions to be modeled and is not characteristic of firm-level claims data.

After estimating the structural parameters of the WYDM model, I quantify the effects of moral hazard by simulating an individual's behavior and outcomes first when health insurance is required, then again when they face the full price of medical care (i.e., no insurance). I find that moral hazard explains 53.1 percent

[3] A deductible is an accumulated medical expenditure threshold that must be reached within an insurance year before the insurer covers any part of the total price of medical care. A MOX level is an accumulated out-of-pocket expenditure threshold at which an individual's share of the total price of additional medical care is zero.

of insured mean annual medical expenditure.[4] The mean effect, however, is heavily influenced by only a few individuals who are very sensitive to insurance coverage. If the distribution of insurance-induced spending is winsorized at the 99th percentile, the percentage of mean annual expenditure explained by moral hazard falls to 46 percent. Furthermore, 54.4 percent of individuals do not alter their behavior when they become insured. The additional medical care consumption induced by insurance acquisition decreases the average number of illnesses acquired over the course of the year by a small but statistically significant amount.

To isolate the effect that within-year dynamics have on the moral hazard estimates, I also estimate a representative annual expenditure (RAE) model. I construct this model by imposing assumptions on the WYDM model so that the incentives faced by an individual and information available to him at the time of his decisions reflect those of an annual expenditure model. That is, both models allow for monthly decision making, but those decisions reflect either the dynamic updating of expectations of the WYDM or retain the annual expectations of the RAE model. The RAE model produces a distribution of insurance-induced spending that is similar in shape to the WYDM model, though the tails of the RAE distribution are heavier. Specifically, the RAE model predicts that 51.9 percent of (insured) mean annual medical expenditure is explained by moral hazard, 42.6 percent of mean expenditure is explained by moral hazard upon winsorizing the distribution at the 99th percentile, and 64.2 percent of individuals do not respond to insurance acquisition; all statistics are significantly different from those produced by the WYDM model. These distributional differences can be explained by differences in the information available to individuals at the time of a medical care decision. Ultimately, the RAE model, which assumes that medical care decisions are made under perfect information, understates the effect that moral hazard has on expenditure by failing to account for an important form of risk protection provided by insurance.

The two models predict insurance-induced spending distributions that differ primarily in the tails, which is important for a number of reasons. According to the Kaiser Family Foundation, 65 percent of the (2010) U.S. medical expenditure was consumed by those in the top decile of the expenditure distribution. Both the WYDM and RAE models suggest that the popularity of health insurance significantly increases the number of extreme spenders; however, the role of insurance is more pronounced for the RAE model. As a result, my findings suggest that annual expenditure models will overstate the effectiveness of insurance-based policies

[4] Analogous simulations generate co-insurance elasticity of medical care demand estimates of 0.20 for the 0-25 co-insurance range and 0.35 for the 25-95 co-insurance range, which compares to elasticity estimates from the RAND Health Insurance Experiment of 0.17 and 0.20, respectively (Manning et al., 1987).

meant to limit high cost elective procedures, such as capitation payments. Another reason to be interested in the full distribution of insurance-induced spending is because the mean can be quite sensitive to behavior in the right tail, which is evident from both WYDM and RAE models. As such, a more conservative estimate of the mean moral hazard effect may limit the relative importance of right tail spenders. Using the (99th percentile) winsorized spending distributions and a series of back-of-the-envelope calculations described in Sections 6, I predict the increase in total U.S. medical expenditure caused by forcing the roughly 41 million uninsured individuals in the U.S. (in 2013) into health insurance coverage. I find the WYDM model prediction to be 11 billion dollars larger than that of the RAE model.

This research contributes to several distinct literatures. First, I estimate insurance-induced moral hazard effects using a structural model of individual-level annual health insurance decisions and within-year medical care consumption. The existing structural literature typically estimates these effects using models of annual medical care decision making (Cardon and Hendel, 2001; Einav et al., 2013; Kowalski, 2015; Bajari et al., 2014). Two recent papers, Einav et al. (2015) and Dalton et al. (2017), also estimate moral hazard effects using structural, within-year decision-making models, but differ from my research in several ways: (i) these models do not allow for endogenous health transitions or endogenous insurance selection and (ii) analysis is limited to prescription drug decisions of Medicare enrollees. Most importantly, neither paper compares estimated moral hazard effects across models featuring both annual and within-year medical care decision making, which is the primary objective of this research.

Second, this paper relates to a large literature focused on estimating insurance-induced moral hazard effects using reduced-form methods (e.g., Manning et al., 1987; Finkelstein et al., 2012; Kolstad and Kowalski, 2012). The structural estimation strategy used in this paper offers several advantages over more reduced-form approaches. For example, unlike structural papers, reduced-form papers focused on estimating the price/co-insurance elasticity of medical care demand must characterize health insurance plans that induce a non-linear price schedule by a single price. Aron-Dine et al. (2013) show that elasticity estimates produced by this research are highly sensitive to this characterization.[5] Another advantage of the structural technique is that it enables the study of a robust set of counterfactual insurance policies, such as those studied in Section 6.5, many of which cannot be studied with a reduced-form model. For example, any regulation that alters the insurance alternative set offered by employers can only be studied in a setting where insurance decisions are modeled

[5]In particular, Aron-Dine et al. (2013) show that the often cited RAND-HIE price elasticity estimate of -0.2 could be as large as -0.6 or as small as -0.04 depending on elasticity definition and price summary method used.

explicitly, as policy implementation may change an individual's insurance decision. Nearly all reduced-form models in this literature use (quasi-)random insurance assignment to identify moral hazard effects, which explicitly prevents insurance decisions from being modeled. A final advantage of the structural approach is that it allows one to study the welfare implications of policy experiments, whereas the reduced-form approach only allows one to study how observed, modeled outcome variables are affected by policy changes.

Third, much like Gilleskie (1998), Khwaja (2010), and Cronin et al. (2017), this paper can be viewed as an empirical application of the influential Grossman (1972) health capital model. A defining feature of this class of models is that medical care is treated as an investment, where the primary motivation for consumption is the improvement of future health. Though the impact of medical care on health transitions is difficult to identify (De Nardi et al., 2016), an additional, more salient, issue is that health status simply cannot be observed independent of medical care consumption in the administrative databases that are popular in the literature. Zweifel and Manning (2000) identify an empirical test for Grossman's model, suggesting that insurance acquisition should result in a sudden increase in expenditure, which ultimately declines after health stock accumulates. The unique combination of insurance decisions, within-year medical care decisions, and endogenous health transitions featured in the WYDM model allows for the first execution of this test in a structural setting. Manning et al. (1985) find support for this hypothesis in the RAND HIE data, but only for dental care. My results show that insurance acquisition also affects doctor's office visits in a way that is consistent with Grossman's health capital model.

Fourth, this research studies the spending and welfare implications of the recent and dramatic increase in high-deductible health plans (HDHP) in the U.S. According to Towers Watson (2016), 77 percent of employers offered a HDHP and 20 percent offered *only* HDHPs in 2016, up from 54 and 4 percent, respectively, in 2010. I conduct two related counterfactual experiments. The first experiment adds a HDHP to every individual's alternative set. I find the addition to be most beneficial for (i) uninsured individuals, who tend to receive less generous insurance offers from their employers, and (ii) relatively healthy insured individuals, who are willing to sacrifice coverage for a lower premium. The second experiment replaces the most generous plan offered to each individual with the same HDHP.[6] Abstracting from supply-side implications, I find that the welfare losses caused by removing the most generous plan are 2.2 times larger than the welfare gains associated with adding the HDHP. The current literature does not speak to the welfare implications of this shift, which is ex-ante

[6]Consistent with national trends, the firms studied by Einav et al. (2013) and Handel and Kolstad (2015) made similar changes to their insurance offerings during and after the study periods, respectively.

ambiguous, as HDHPs are characterized by greater cost-sharing *and* lower premiums - a trade-off that favors low-risk individuals. A small literature does examine the impact of HDHPs on spending, utilizing claims data from both the firms (e.g., Brot-Goldberg et al., 2017) and insurers (e.g., Lo Sasso et al., 2010). As a result, none of the existing literature speaks to how the uninsured have responded to the shift towards HDHPs.

The paper is organized as follows. Section 2 provides motivation for research on moral hazard effects and the WYDM model. Section 3 details the theoretical model of insurance and medical care demand. Section 4 describes the data used in estimation. Section 5 details the estimation procedure and discusses identification. Section 6 contains all empirical results. Section 7 concludes.

2 Motivation and Background

Health insurance generates welfare by protecting a risk averse individual from medical expenses associated with unforeseen health shocks (Arrow, 1963). The welfare gains from risk protection are potentially mitigated by changes in individual behavior after becoming insured. For example, insurance lowers the out-of-pocket price of medical care, which can lead to excess consumption when sick, known as ex-post moral hazard (Pauly, 1968). Also, a reduction in the expected cost of curative medical care can reduce participation in healthy behaviors (e.g., preventative medical care) leading to worse health outcomes and potentially greater medical expenditure in the future, known as ex-ante moral hazard. Each of these behavioral responses affects how much, if any, welfare is generated by insurance coverage (Cutler and Zeckhauser, 2000). Therefore, efficient health insurance plan design requires an understanding of how health insurance leads to changes in individual medical care consumption behavior.

The primary challenge in estimating the change in medical expenditure that is caused by health insurance possession and/or generosity is the endogenous selection of health insurance. If those who expect to consume more medical care during a health insurance year select generous health insurance coverage, known as adverse selection (Akerlof, 1970), then the positive correlation observed between medical expenditure and insurance generosity could be due to either moral hazard or selection effects.[7] One method that has been used to separate these effects is a randomized experiment, which eliminates the endogenous selection

[7]Whether this correlation is driven by moral hazard or adverse selection has important policy implications. The policy prescription for moral hazard is less risk protection through greater cost-sharing, reducing an individual's incentive to overconsume care. The negative implications of adverse selection are reduced by risk pooling.

problem by randomizing insurance possession. A well known example is the 1971 RAND Health Insurance Experiment (HIE), which randomly distributed health insurance plans to participants in six U.S. cities and collected health and medical care consumption data in the years following (Newhouse and Group, 1993). A more recent example is the 2008 Oregon HIE, which expanded the state's Medicaid program to 10,000 additional low-income adults via lottery (Finkelstein et al., 2012). There are also numerous quasi-experimental studies that have leveraged plausibly exogenous changes in public insurance programs, such as Medicaid expansion (e.g., Dafny and Gruber, 2005) or the Massachusetts market reforms (e.g., Kolstad and Kowalski, 2012), to separately identify moral hazard effects from selection effects.

This experimental and quasi-experimental research is primarily focused on measuring the change in outcome variables caused by the introduction of a specific public policy, which makes the results from this literature difficult to generalize. In response, a number of researchers have designed and estimated structural models of health insurance and subsequent medical care decisions. Thus far, most of these models that have aggregated medical expenditures and health outcomes up to the annual (Cardon and Hendel, 2001; Einav et al., 2013; Handel, 2013; Kowalski, 2015; Bajari et al., 2014; Handel and Kolstad, 2015) or biennial (Khwaja, 2010) level. Annual expenditure models have been popular, partly due to computational convenience, as within-year decision-making models require specifying and solving a DPP. Also, annual expenditure models can be estimated without high frequency health or illness data, which is both difficult to find and necessary when estimating a model of within-year medical care decisions.[8] This research builds on these structural annual expenditure models by allowing for monthly medical care decisions to be made over a health insurance year and by relaxing several common assumptions.[9] Section 1 of the web appendix discusses how these differences could lead the models to produce different moral hazard estimates.

[8]Einav et al. (2013), Kowalski (2015), Bajari et al. (2014), Handel (2013), and Handel and Kolstad (2015) use administrative claims data that allow high-frequency medical care decisions to be observed; however, health/illness is only observed when an individual consumes care, preventing endogenous health transitions from being modeled.

[9]Several researchers have studied within-year medical care demand, including Keeler and Rolph (1988), Ellis (1986), Gilleskie (1998), Aron-Dine et al. (2015), Einav et al. (2015), Dalton et al. (2017), and Lin and Sacks (2017). In all of these studies, insurance possession is either randomized or assumed to be exogenously determined, limiting the scope of policy experiments that can be studied.

3 Model

This section describes the optimization problem of an unmarried, childless, non-pregnant, employed individual who makes an annual health insurance decision followed by a sequence of medical care decisions to maximize the value of his expected discounted future utility.[10] The timing of the model can be observed in Figure 1. At the beginning of each year, y, a forward-looking individual observes the set of health insurance alternatives offered by his employer and the presence of any illnesses. Before the start of the first month, $t=1$, he selects the health insurance alternative that maximizes his expected discounted future utility. Among other things, this expected utility is a function of anticipated medical care behavior within the year conditional on insurance coverage.

At the beginning of each month, an individual learns his illness state, which evolves stochastically over the course of the year and is influenced by illness history and previous medical care consumption. After learning his current illness state, the individual decides how much and what types of medical care to consume. The amount he pays for a unit of medical care depends on the unit price, the cost-sharing characteristics of his health insurance plan, and his accumulated medical expenditure within the coverage year. Much like the price uncertainty an individual faces in the U.S. medical care market, the total price of care is stochastic over time and unknown prior to consumption. Conditional on consuming medical care, the individual learns the price he paid and updates his stock of accumulated out-of-pocket expenditure before transitioning into the next month. The remainder of this section explains the model and solution in greater detail.

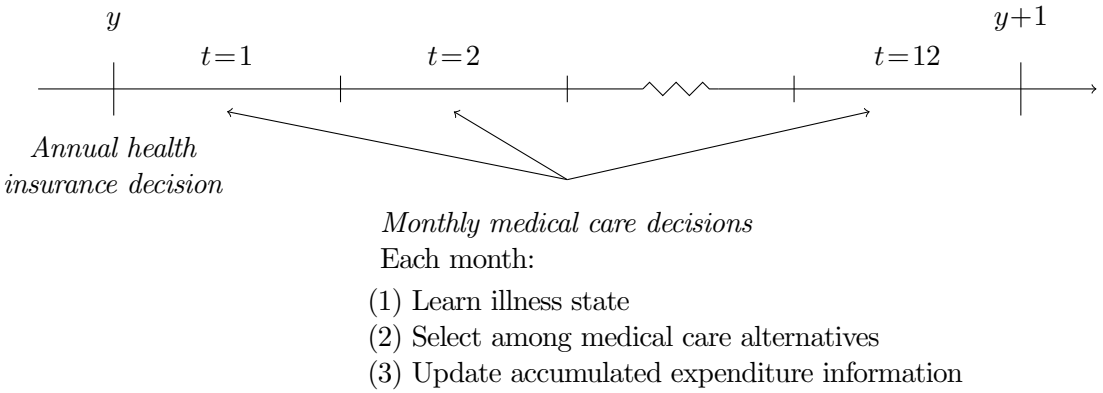
3.1 Annual and Monthly Decisions

At the beginning of each year, an individual observes the set of health insurance plans offered to him by his employer. Each plan is defined by its out-of-pocket premium, network type, and a set of cost-sharing characteristics that enter an individual's budget constraint throughout the year, determining how much is paid out of pocket for medical care.[11] I define an indicator function, I_{iy}^{j}, that equals one if individual

[10] The model focuses on individuals receiving an employer-sponsored health insurance (ESHI) offer because insurance information is only available for these individuals in the data. Of the non-elderly, adult population in 2014: 55% held ESHI, 12% were uninsured, 26% were insured by state or federal governments, and 7% were privately insured (Kaiser Family Foundation). The model focuses on a single, childless individual in order to capture dynamic health, non-linear out-of-pocket prices, and medical care demand throughout the insurance coverage period, which would be computationally prohibitive if it involved more than one individual.

[11] Possible cost-sharing characteristics include a deductible, MOX level, co-insurance rate, and co-pay level. A deductible and MOX level are defined in footnote 3. A co-insurance rate is the share of the medical care price that an individual must

Figure 1: Timing of the Model



i selects insurance plan j in year y and zero otherwise.[12] Only one plan can be held at a time, so that $\sum_{j\in J^i} I^j_{iy}=1\forall i$ and y, where J^i is individual i's set of ESHI plans, including the option to decline coverage.

In each month, an individual learns his illness state (defined below) before making a medical care decision. He chooses the number of doctor visits, d_{it}; hospital days, h_{it}; and whether or not to consume prescription drugs, r_{it}.[13] The monthly medical care decision is represented by an indicator function, m^{dhr}_{it}, that equals one if an individual chooses the bundle $\{d,h,r\}$ and zero otherwise. Bundles are mutually exclusive with a maximum of D doctor visits and H hospital days in each month, such that $\sum_{d=0}^{D}\sum_{h=0}^{H}\sum_{r=0}^{1} m^{dhr}_{it}=1\forall i$ and t.

3.2 Illness Transitions and Probabilities

An individual's illness state evolves stochastically over the course of the insurance year. The illness state in month t is defined by the number of acute, A_{it}, and chronic, C_{it}, illnesses.[14]

I define an acute illness as any medical condition that eventually subsides and, under normal conditions, has

pay out of pocket; the remainder is paid by the insurer. A co-pay level is a fixed dollar amount that an individual must pay out of pocket for a unit of medical care; again, the remainder is paid by the insurer. Note that deductibles, co-insurance rates, and co-pay levels frequently vary across prescription drugs, doctor's office visits, and hospital visits in the data.

[12]For notational simplicity and consistency, I include the subscript i to describe individual-level variables only when defining the variable. The subscript i is suppressed thereafter.

[13]According to the Centers for Disease Control and Prevention (CDC) these three types of care account for over 80 percent of personal medical expenditures in the US. For the population of study, the percentage is even higher because non-elderly individuals are unlikely to consume nursing home or home health care. Other relevant medical care products, such as dental and optical, are unlikely to be covered by standard ESHI plans and are thus excluded.

[14]Death is only observed once in the data because the estimation sample includes ages 19-64, so it is not modeled as a possible health outcome. A *death state* would be a simple addition with alternative data.

no permanent effect on an individual's health or medical care consumption. This characterization describes both short-natured ailments, such as a common cold or influenza, as well as persistent but non-permanent conditions, such as a pneumonia or a broken bone. In estimation, the number of acute illnesses that an individual has in month $t+1$ is determined by the latent variable A^*_{t+1}, which can be written as

$$A^*_{t+1} = G_A(\mathbf{W_{t+1}}, A_t, C_t, m_t^{dhr}; \alpha) + \mu_1^k + \psi_{t+1} \tag{1}$$

where $\mathbf{W_{t+1}}$ is a vector of individual characteristics such as sex, race, income, education, MSA indicator, age, and calendar month indicators; α is a vector of parameters; μ_1^k captures permanent unobserved heterogeneity for an individual of type k (see Section 5.2 for details); and ψ_{t+1} is an i.i.d error. The function $G_A(\cdot)$ is linear in parameters and includes both main and interaction effects, which allows the productivity/efficacy of medical care to vary by illness state entering the month.

I assume that A_t follows an ordered structure, so that $A_t = q$ if $\kappa^a_q < A^*_t \leq \kappa^a_{q+1}, \forall q \in \{0,...,4\}$. Define $\widetilde{A}^*_t = A^*_t - \psi_t$ and $\Lambda(\cdot)$ as the logistic function. Assuming ψ_t follows a logistic distribution, the (ordered logit) probability of having q acute illnesses is[15]

$$P(A_t = q) = \pi_t^q = \Lambda(\kappa^a_{q+1} - \widetilde{A}^*_t) - \Lambda(\kappa^a_q - \widetilde{A}^*_t). \tag{2}$$

I define a chronic illness to be any medical condition that never subsides (e.g., diabetes, asthma) or, under normal conditions, has a permanent effect on an individual's health or medical care consumption (e.g., cancer, stroke, hypertension). As such, the occurrence of a chronic illness is modeled as a permanent, absorbing state. Medical care is then used to prevent or control a chronic illness, lessening the impact that one chronic illness has on the development of another chronic or acute illness. Because chronic illnesses are permanent, the model determines how many *additional* chronic illnesses an individual acquires in each month, which I define as C^+_t. In estimation, additional chronic illnesses in month $t+1$ are determined by the latent variable C^{+*}_{t+1}, which can be written

$$C^{+*}_{t+1} = G_C(\mathbf{W_{t+1}}, A_t, C_t, m_t^{dhr}; \delta) + \mu_2^k + \zeta_{t+1} \tag{3}$$

As with acute illnesses, I assume C^+_t follows an ordered logit structure. The probability of acquiring q additional chronic illnesses in month t is written as $\gamma_t^q, \forall q \in \{0,1,2\}$.

[15] $\{\alpha, \mu_1^k, \kappa_2^a, \kappa_3^a, \kappa_4^a\}$ are parameters to be estimated. $\{\kappa_0^a, \kappa_1^a, \kappa_5^a\}$ are normalized to $\{-\infty, 0, \infty\}$.

3.3 Utility Function and Budget Constraint

Preferences for medical care consumption bundle $\{d,h,r\}$ in month t are described by the following contemporaneous utility function

$$U(m_t^{dhr},\mathbf{S_t}|\mathbf{p_t}) = \frac{X_t^{\omega_0 \mathbf{R}}}{\omega_0 \mathbf{R}} + G_U(\mathbf{W_t},A_t,C_t,m_t^{dhr};\omega) + \mu(m_t^{dhr},k) + \epsilon_t^{dhr} \tag{4}$$

where $\mathbf{S_t} = (\mathbf{W_t},A_t,C_t,ADE_t,AHE_t,I_y^j,\mu^k,\epsilon_t^{dhr})$ is the information known by an individual at the time of his medical care decision (i.e., his *state*); $\mathbf{p_t} = (p_t^d,p_t^h,p_t^r)$ is a vector of doctor, hospital, and prescription drug prices; X_{it} is consumption of non-medical goods, which is defined in Equation 5; $\mathbf{R} = (1,sex,race,age)$; ω is a vector of linear parameters; $\mu(m_t^{dhr},k)$ captures permanent unobserved heterogeneity in medical care preferences for an individual of type k; and ϵ_{it}^{dhr} is the unobserved utility received from medical care bundle $\{d,h,r\}$.[16] For simplicity, Equation 4 can be rewritten as the sum of deterministic and random components, $\overline{U}(m_t^{dhr}) + \epsilon_t^{dhr}$.

The monthly budget constraint is

$$X_t = Y_t - P_{jt} - O(m_t^{dhr},\mathbf{p_t},ADE_t,AHE_t,I_y^j) \tag{5}$$

where Y_{it} is monthly income; P_{ijt} is the month t plan j out-of-pocket premium; $O(\cdot)$ is the out-of-pocket expenditure on medical care in month t; and ADE_{it} and AHE_{it} represent accumulated out-of-pocket medical expenditure for doctor visits and hospital days entering month t, respectively.[17]

3.4 Medical Care Prices and Expenditure

Two characteristics of the U.S. medical care market make the within-year medical care decision-making problem an important economic construct. First, most individuals do not pay the total price of medical care because of a cost-sharing arrangement with their health insurance provider. Rather, individuals pay an out-of-pocket price that is determined by the total price, insurance plan characteristics, and accumulated medical expenditure during the coverage year. For example: an individual with a $300 deductible, 10 percent

[16] I allow for permanent unobserved heterogeneity in the utility from medical care consumption by parameterizing $\mu(\cdot)$ such that $\mu(m_t^{dhr},k) = d\mu_{3a}^k + d^2\mu_{3b}^k + h\mu_{4a}^k + h^2\mu_{4b}^k + r\mu_5^k + (1-m_{it}^{000})\mu_6^k$.

[17] This structure assumes that all income is consumed by the end of each month, as monthly savings are not observed in the data. In their analysis of health insurance, savings, and retirement behavior, French and Jones (2011) explain that omitting savings from an individual's dynamic problem ignores the ability to smooth consumption through savings, which can overstate the value of insurance. However, omitting savings from their model has no impact on retirement decisions, suggesting that omitting savings from this model may not impact insurance decisions.

co-insurance rate, and \$0 of accumulated expenditure who is charged \$100 for a doctor visit pays the full \$100 out of pocket. However, if the same individual were to have accumulated \$250 in medical expenditure prior to the visit, then he would pay only \$55 out of pocket. An individual with health insurance characterized by such a cost-sharing structure faces a non-linear budget constraint. The out-of-pocket expenditure function, $O(\cdot)$, is constructed so that the budget constraint in Equation 5 contains these non-linearities. The precise calculation of out-of-pocket expenditure is detailed in Section 2 of the web appendix.

A second unique characteristic of the U.S. medical care market is that individuals are typically uncertain of the total price of medical care prior to consumption (Painter and Chernew, 2012; Rosenthal et al., 2013). A lack of menu prices, uncertainty of diagnosis, and wide price variation even in local markets contribute this uncertainty. To address the reality of uncertain prices, I assume that an individual does not observe total medical care prices prior to making a monthly medical care decision. Rather, an individual knows the conditional distributions from which doctor, hospital, and prescription drug prices are drawn. An individual makes medical care decisions by integrating over these conditional price distributions, which are estimated from the data.

Total price distributions are defined as $f^d(p_t^d|\mathbf{S_t};\lambda^d)$, $f^h(p_t^h|\mathbf{S_t};\lambda^h)$, and $f^r(p_t^r|\mathbf{S_t};\lambda^r)$, where $\{\lambda^d,\lambda^h,\lambda^r\}$ are parameters to be estimated. An individual's state, $\mathbf{S_t}$, contains the variables HMO_j, PPO_j, and FFS_j, which are indicators of the plan's coverage type.[18] Controlling for coverage type captures the lower rates negotiated by insurance providers who contract with a network of physicians. The MSA indicator in $\mathbf{S_t}$ should capture urban area variation in prices. These distributions also depend on individual observed illness states and initial self-reported health status. Finally, medical care price shocks are likely to be correlated with one another, as well as the unobserved illness shocks; an individual who receives an exceptionally bad illness shock is also likely to face price distributions that are shifted rightward or have fatter right tails. To control for these potential sources of unobserved correlation, I allow the permanent unobservables that influence preferences, illness states, and initial conditions (i.e., μ^k) to also influence the price distributions.

[18] A health maintenance organization (HMO) limits its enrollees to receiving medical care from a specified group of providers. A preferred provider organization (PPO) defines a preferred network of providers from which care can be purchased less expensively. Fee-for-service (FFS) plans cover an enrollee equally at all medical care providers. The model does not differentiate between in- and out-of-network medical care consumption; all is assumed to be in-network and all cost-sharing characteristics are specific to in-network consumption.

3.5 The Optimization Problem

An individual maximizes his expected discounted future utility by selecting the optimal sequence of medical care bundles, m_t^{dhr}, for $t=1,...,T$ and insurance plans, I_y^j, for $y=1,...,Y$ conditional on his state variables, $\mathbf{S_t}$. The dynamic optimization problem has two stages, as insurance decisions are made annually and medical care decisions are made repeatedly over the course of a year.

3.5.1 The Optimal Monthly Decision Rule

Let $V_{dhr}^{ac}(\cdot_t)$ be the month t value of expected discounted future utility for medical care decision m_t^{dhr} and illness state $(A_t=a, C_t=c)$. Using Bellman's (1957) Equation, this value is constructed as the sum of the contemporaneous and expected discounted future utility yielded by the alternative. Conditional on unobserved heterogeneity type k, insurance plan j, and medical care price vector $\mathbf{p_t}$, the alternative-specific value function is written as

$$V_{dhr}^{ac}(\mathbf{S_t},\epsilon_t^{dhr}|\mu^k,I_y^j,\mathbf{p_t}) = \overline{U}(m_t^{dhr}) + \epsilon_t^{dhr} + \beta\left[\sum_{a'=0}^{4}\pi_{t+1}^{a'}(\mathbf{S_t},m_t^{dhr})\sum_{c'=0}^{2}\gamma_{t+1}^{c'}(\mathbf{S_t},m_t^{dhr})\left[V^{a',c+c'}(\mathbf{S_{t+1}}|\mu^k,I_y^j)\right]\right] \tag{6}$$

for $t<T$, and for $t=T$

$$V_{dhr}^{ac}(\mathbf{S_t},\epsilon_t^{dhr}|\mu^k,I_y^j,\mathbf{p_t}) = \overline{U}(m_t^{dhr}) + \epsilon_t^{dhr} + \beta Q_{y+1}(\mathbf{S_0}) \tag{7}$$

where β is the monthly discount factor and $Q_{y+1}(\mathbf{S_0})$ is the value of expected discounted future utility in month $t=0$ of year $y+1$, (i.e., prior to the year $y+1$ insurance decision). Maximal expected utility in illness state $(A_{t+1}=a, C_{t+1}=c)$ in month $t+1$ is

$$V^{ac}(\mathbf{S_{t+1}}|\mu^k,I_y^j) = E_t\left[\max_{dhr} V_{dhr}^{ac}(\mathbf{S_{t+1}},\epsilon_{t+1}^{dhr}|\mu^k,I_y^j)\right]. \tag{8}$$

The expectation operator is subscripted by t because an individual must form this expectation prior to learning month $t+1$ medical care preference shocks, ϵ_{t+1}^{dhr}.

The value function in Equation 6 conditions on realized medical care prices; however, I assume that an individual only knows $f^d(\cdot)$, $f^h(\cdot)$, and $f^r(\cdot)$, the conditional distributions from which prices are drawn. Solution to the optimization problem requires integration over these price distributions. Defining $f^*(\mathbf{p_t}) = f^d(p_t^d) * f^h(p_t^h) * f^r(p_t^r)$, the value function written unconditional on prices is

$$V_{dhr}^{ac}(\mathbf{S_t},\epsilon_t^{dhr}|\mu^k,I_y^j) = \int_{\mathbb{R}_+^3} f^*(\mathbf{p_t}) V_{dhr}^{ac}(\mathbf{S_t},\epsilon_t^{dhr}|\mu^k,I_y^j,\mathbf{p_t}) d\mathbf{p_t} \tag{9}$$

Conditional on the prior insurance decision and unobserved heterogeneity, a utility maximizing individual selects medical care consumption bundle $\{d,h,r\}$ with probability

$$P(m_t^{dhr}=1)=P\Big[V_{dhr}^{ac}(\mathbf{S_t},\epsilon_t^{dhr}|\mu^k,I_y^j)\geq V_{d'h'r'}^{ac}(\mathbf{S_t},\epsilon_t^{d'h'r'}|\mu^k,I_y^j)\quad\forall\quad d'h'r'\Big]. \tag{10}$$

3.5.2 The Optimal Annual Decision Rule

The DPP can be solved backwards to recover the time $t=0$, year y value function conditional on a chosen health insurance alternative $j\in J_y^i$. That is,

$$V(\mathbf{S_0}|\mu^k,I_y^j)=\sum_{a'=0}^{4}\pi_1^{a'}(\mathbf{S_0},m_{T,y-1}^{dhr})\sum_{c'=0}^{2}\gamma_1^{c'}(\mathbf{S_0},m_{T,y-1}^{dhr})\Big[V^{a',c+c'}(\mathbf{S_1}|\mu^k,I_y^j)\Big] \tag{11}$$

where $m_{T,y-1}^{dhr}$ is the medical care decision in the last month of year $y-1$ and c is the number of chronic illnesses possessed entering year y. Stated explicitly, Equation 11 represents the discounted value of optimal future behavior calculated at the beginning of year y unconditional on the first month illness state but conditional on insurance plan j (i.e., the expected discounted future value of plan j). I assume that this value, plus an additively separable i.i.d random error, determines the discounted expected future value of plan j,

$$Q_y^j(\mathbf{S_0},\phi_y^j|\mu^k)=V(\mathbf{S_0}|\mu^k,I_y^j)+\phi_y^j. \tag{12}$$

Thus, a utility maximizing individual selects insurance plan j with the probability

$$P(I_y^j=1)=P\Big[Q_y^j(\mathbf{S_0},\phi_y^j|\mu^k)\geq Q_y^{j'}(\mathbf{S_0},\phi_y^{j'}|\mu^k)\quad\forall j'\Big]. \tag{13}$$

Several papers in the health insurance choice literature have argued against the use of an i.i.d random error (Einav et al., 2013; Handel, 2013; Kowalski, 2015). Importantly, each of these papers uses data from a single firm where plans are financially rankable and identical on non-financial characteristics. In such a circumstance, (i) it is not clear what an error component would capture that is not modeled and (ii) the i.i.d assumption is likely inappropriate given the financial rankability of plans. MEPS participants work for hundreds of firms offering many different insurance alternative sets; therefore, I cannot verify that plans within a firm are vertically ranked and there is little evidence to suggest that this is a good assumption generally. An i.i.d. random error is not without precedence in the literature (Cardon and Hendel, 2001; Khwaja, 2010). A plan-specific error allows the model to rationalize observed insurance decisions using both modeled and unmodeled factors. Absent this error, the model must rationalize all observed decisions, potentially tasking structural parameters with explaining behavior outside of the model. For example,

Handel and Kolstad (2015) suggest that inertia and mis-information about plan features are important determinants of health insurance choice; however, neither of these factors are controlled for in the WYDM model. The assumption that errors are independent also has the benefit that it smooths out the likelihood function, which is computationally attractive for numerical optimization.

4 Data

4.1 Description of MEPS

My empirical analysis uses data from the 1996-1999 cohorts of the Medical Expenditure Panel Survey (MEPS), which is collected by the Agency for Healthcare Research and Quality (AHRQ). The MEPS contains detailed health, medical expenditure, health insurance, and demographic information for a nationally representative sample of households in the U.S. New participants are added annually, drawn randomly from the previous year's National Health Interview Survey sample. Each cohort is interviewed five times over the two years that follow January 1st of the cohort year.

The MEPS has two features that make it particularly well suited for the purposes of this research. First, detailed employer-level insurance information exists for the individuals comprising the 1996-1999 cohorts. AHRQ used information gathered in the first interview to contact current main employers, from which they obtained premium and cost-sharing characteristics for all plans offered to the employee. This data feature, which is unique in national survey data, enables me to model a health insurance decision from the full set of available alternatives for an individual with a participating employer. Second, illness information in the MEPS file is initially self-reported. An individual is asked to describe any, "health problems that may have bothered (person)," over the survey period, independent of whether they sought medical care. Professional coders convert these descriptions into three-digit ICD-9-CM codes, some of which are later verified by medical professionals if care is sought. The ability to observe illness episodes even when medical care is not consumed is an important feature of these data that is not characteristic of firm-level claims data, as it allows endogenous illness transitions to be modeled. Moreover, note that 34 percent of the illnesses reported by the estimation sample do not result in any type of medical care consumption, suggesting that these illnesses represent a significant source of unobserved heterogeneity in studies that utilize firm-level claims data. That said, self-reported illness data have the potential to introduce measurement error to the model. Of particular concern is the possibility that those with health insurance are more likely to visit a health professional,

receive a diagnosis, and, thus, report an illness. The likelihood of such endogenous illness reporting and the implication for my empirical results are discussed extensively in Section 3.5 of the web appendix.

A number of important assumptions are required to prepare the data for estimation. For example, ICD-9-CM medical codes must be interpreted as an acute or chronic illness. Also, medical care consumption dates and partially observable illness dates must be used to determine the starting and ending month of illnesses reported five times over the course of two years. I also face the challenge that at least one insurance cost-sharing feature is missing in about half of the observed plans; imputations were made in these situations using a number of different techniques. The magnitude of these complications, and others, and the assumptions required to overcome them are discussed at length in Section 3 of the web appendix.

4.2 Determination of the Sample

The sample used in estimation is taken from the nationally representative sample of single and childless individuals included in the 1996-1999 cohorts of the MEPS. I focus on employed individuals between the ages of 19 and 64 whose employers sponsor health insurance coverage.[19] I exclude the unemployed and those employed without an insurance offer because limited insurance information was gathered for these individuals. Employed individuals who receive an insurance offer but choose to be uninsured *are* included in the estimation sample. These omissions are representative of restrictions found in similar work.

Sample inclusion requires that an individual's offered ESHI plans are observed in the employer link file. Roughly 50 percent of MEPS participants lack these data due to employee and/or employer refusal to disclose the information. Individuals must also participate in all interviews during the insurance year. Finally, I restrict the sample to those either: (i) holding ESHI, and no outside coverage, for an entire year; or (ii) remaining completely uninsured all year. I do not model within-year insurance switching and cannot observe privately purchased or government provided (e.g., Medicaid) plan characteristics; therefore, individuals switching plans or holding outside coverage are dropped. See Section 3.3 of the web appendix for sample size by inclusion criteria.

The final estimation sample contains 1,232 individuals or 14,784 person-month observations. In Section 4 of the web appendix, I compare the estimation sample to that of a nationally representative sample of 19-64 year olds, who are employed with an ESHI offer - a population for which one may wish to draw inference. The samples are very similar, though the estimation sample is comprised of more government employees

[19] I also exclude full time students under the age of 24 because 1996-1999 federal law allowed these individuals access to their parent's health insurance. Individuals over 64 are excluded because they have access to Medicare.

and employees of large companies, which is expected, as these employers are the most likely to respond to the survey.[20] In comparison, the existing literature's reliance on claims data from single private firms results in estimation samples that are uniquely specific to industry and geography, which might impact the externally validity of research findings. For example, Einav et al. (2013) use a sample of unionized, hourly-wage employees of Alcoa, Inc. Individuals in this sample are more likely to be male, non-white, and low income than a similarly defined subsample of the nationally representative MEPS.

4.3 Sample Statistics

Table 1 compares insured and uninsured individuals in the estimation sample. The insured are older, more educated, wealthier, and are more likely to be white and female. The insured are also more likely to enter the year with a chronic illness, are more likely to get an acute illness during the year, and experience more months with an acute illness; providing some evidence of adverse selection. The insured also consume more units and greater values of doctor and prescription drug care. The percentage of the population that consumes at least one hospital day during the year seems large (25 percent of insured and 17 percent of uninsured individuals), but includes emergency room visits as well as outpatient and inpatient visits. The uninsured are less likely to have at least one hospital day, but on average consume a larger number of hospital days than the insured, which is due to higher emergency room usage among the uninsured. To reduce estimation time, I limit the maximum number of doctor visits and hospital days in a month to nine and five, respectively.[21] The insured face lower prices for doctor visits and hospital days and higher prices for prescription drugs.[22] The large variance in prices is due

[20] As a result of this difference, the estimation sample is also likely to contain individuals with access to a larger, more generous set of health insurance alternatives. I discuss how these small differences between the samples may impact the external validity of my results in Section 4 of the web appendix.

[21] Monthly doctor visits (hospital days) exceeds the maximum of nine (five) only 33 (30) times. In these instances, the number of visits/days is set to the maximum and the average price paid for a unit of medical care is adjusted accordingly. For example, if an individual visits the doctor 12 times in a month with an average price of $100, then the data are adjusted so that he visits the doctor nine times with an average price of $133.

[22] Finding that the insured pay higher monthly prescription drug prices than the uninsured is an expected by-product of simplifying assumptions that are imposed on the data. The model does not allow for selection on drug quality and, within a month, only allows for a drug decision on the extensive margin. If the average uninsured individual is more likely to purchase generic medications and/or fills fewer prescriptions in a month conditional on filling any, then average (monthly) prices will be lower than that of an insured individual. The consequence of this simplifying assumption is that the model may overstate the benefit of prescription drug consumption for the uninsured; however, I find little evidence that this has a significant effect on moral hazard estimates.

Table 1: Sample Statistics by Insurance Status

	Insured		Uninsured	
	Mean	S.D.	Mean	S.D.
Demographics (time invariant)				
age	40.04	11.62	34.73	11.44
education (highest grade completed)	13.73	2.43	12.46	2.49
income (in 1996 dollars)	35198.52	22120.24	19011.55	13144.27
male	0.48		0.52	
lives in a MSA	0.82		0.83	
Hispanic	0.11		0.21	
black	0.14		0.19	
federal employee	0.09		0.09	
Initial Self-Reported Health (time invariant)				
excellent	0.35		0.36	
very good	0.34		0.32	
good	0.23		0.22	
fair	0.07		0.09	
poor	0.01		0.01	
Illness (time varying)				
entered year with chronic illness	0.40		0.24	
chronic illness by years end	0.50		0.32	
average number of chronic illnesses†	1.84	1.24	1.79	1.18
at least one acute illness during sample year	0.82		0.74	
total months with acute illness	5.60	4.52	3.04	3.21
average number of acute illnesses†	1.56	0.90	1.31	0.67
Medical Care Prices (time varying)				
transaction price for a doctor visit	89.57	159.09	112.43	160.52
transaction price for a hospital day	823.72	1357.51	878.72	2127.58
transaction price for a Rx month	75.47	115.57	51.00	54.60
Medical Care Consumption (time varying)				
at least one doctor visit in sample year	0.74		0.50	
total doctor visits in sample year	5.13	8.95	2.51	4.89
at least one hospital day in sample year	0.25		0.17	
total hospital days in sample year	0.63	1.73	0.81	3.76
at least one Rx month in sample year	0.66		0.45	
total Rx months in sample year	4.69	5.00	2.19	3.87
consumed any preventative care	0.17		0.14	
probability of consumption in ill month	0.59		0.41	
probability of consumption in well month	0.12		0.03	
total annual expenditure on doctor visits	455.72	951.61	289.67	759.21
total annual expenditure on hospital days*	660.46	2706.30	1108.39	6320.84
total annual expenditure on Rx consumption	353.28	819.20	111.48	275.49
Other				
number of offered plans	4.41	5.99	3.05	4.46
Sample				
individuals	1119		113	
person-month observations	13428		1356	

Notes: Income is calculated as the sum of post-tax income, sale earnings, and tax refund. Medical care prices are only observed when medical care is consumed. An individual has *consumed any preventative care* if, during the insurance year, they consumed any form of medical care in a month in which they had no acute or chronic illnesses. In accordance with the model, the prescription drug transaction price is the price of one month of drug consumption. In practice, it is calculated at an individuals total expenditure on prescription drugs in a month.

† Conditional on having any acute/chronic illness.

* One uninsured individual had hospital expenditures totaling $52,032.16. Removing this outlier lowers the mean to $647.93 for the uninsured.

to the coarse classification of medical care types. High priced procedures (e.g., hip replacement surgery) and low priced procedures (e.g., ER visit for a sprained ankle) contribute to the same hospital price distribution.

Chosen and rejected plans are compared in Table A6 of the web appendix. The table suggests that individuals have a general preference for lower premiums and therefore less generous plans. Compared to the average rejected plan, held plans are more likely to have a deductible, are less likely to have a MOX level, set higher thresholds when the plan has a deductible or MOX level, and feature higher co-insurance rates and co-pay levels for both doctor and hospital care. The survey does not elicit cost-sharing information for prescription drugs; therefore, I assume that HMO, PPO, and FFS plans have prescription co-insurance rates of 13%, 17%, and 19%, respectively, and that prescription drug expenditure is completely unrelated to a plan's deductible and MOX level. These assumptions are justified in Section 3.4.2 of the web appendix.

Many of the WYDM model's parameters are identified by individual-level month-to-month variation in state variables. For example, estimates of the productivity/efficacy of medical care are identified by the month-to-month covariance between observed medical care consumption and illness transitions. As such, transition matrices showing substantial within-individual, month-to-month variation in illness and medical care use are presented in Section 3.6 of the web appendix. In Section 3.7 of the web appendix, I highlight the month-to-month variation in effective medical care prices by showing how the average proportion of medical care prices paid out of pocket changes over the insurance year. Among individuals experiencing a change, the average proportion of prices paid out of pocket falls from 97% in the first month of the insurance year to 13% by the final month. Section 3.8 of the web appendix contains results from several reduced-form regressions of medical care demand on marginal and expected end-of-year prices. Unlike the WYDM model, these regressions do not control for price endogeneity, but do supply evidence of significant negative correlation between prices and demand.

5 Empirical Implementation

For each individual in the estimation sample, I observe one health insurance decision followed by medical care consumption, prices, and illness states over the following year. Therefore, I estimate the structural parameters of the model described in Section 3 using one year of data for each individual. In what follows, I discuss estimation challenges, unobserved heterogeneity, and identification. In Section 5 of the web appendix, I construct the likelihood function, in addition to discussing the handling of endogenous initial conditions and the nested fixed point algorithm used in estimation (Rust, 1987).

5.1 Approximating the Future Value of a Medical Care Alternative

Solving the optimization problem requires calculating an individual's value function for each medical care bundle in each month. According to Equation 6, the value of a bundle in month t, $V_{dhr}(\mathbf{S_t})$, is a function of the maximal expected utility in the next month, $V(\mathbf{S_{t+1}})$, where the future state vector, $\mathbf{S_{t+1}}$, is unknown. Thus, in order to calculate $V_{dhr}(\mathbf{S_t})$ in practice, a value $V(\mathbf{S_{t+1}})$ is needed for every potential outcome of $\mathbf{S_{t+1}}$ following every potential history of outcomes ($\mathbf{S_0}$,...,$\mathbf{S_t}$). Given the number of months in a year, the number of variables in the state vector, and the fact that several of the state variables are continuous, the number of required future values grows exponentially.

To avoid what Bellman (1957) refers to as the "curse of dimensionality," I use an interpolation technique developed by Keane and Wolpin (1994) to approximate an individual's maximal expected future value in each month. The method works as follows: beginning in the last month of a year, $t=T$, I draw 3,500 random outcomes of the state vector, $\mathbf{S_T}$. I then calculate maximal expected future utility for each draw according to Equation 8. By estimating a linear regression of these values on the state variables in $\mathbf{S_T}$, I generate a mapping from any possible state to expected future values.[23] This mapping can be used in month $T-1$ to approximate the maximal expected future value of month T, because each alternative in month $T-1$ generates a probability distribution over $\mathbf{S_T}$. By repeating this process backwards, I can solve the model back to month $t=0$.

A related challenge in solving any finite horizon dynamic problem is determining the maximal expected future value in the final period; $Q_{y+1}(\mathbf{S_0})$ from Equation 7. I take the popular approach of formulating a terminal value function. The value is determined by a non-stochastic linear function of the state variables entering the first month of the following year, medical care consumed in month T, and a vector of parameters that are estimated as part of the MLE procedure.

5.2 Unobserved Heterogeneity

The estimation procedure assumes that each individual can be described by a permanent, unobserved type, which allows for correlation between the unobserved determinants of choices and outcomes in the model. The strategy decomposes the model's error terms into two additively separable components: an i.i.d.

[23] This regression includes as controls (i) state variables, which includes insurance plan characteristics; (ii) interactions between state variables; and (iii) the deterministic components of the current month utility function, which is advised by Keane and Wolpin (1994). These results are available upon request.

serially-uncorrelated random error (e.g., ψ_{t+1} in Equation 1) and μ^k, a persistent unobserved component that varies across individuals of $k=1,...,K$ different types. I assume that the distribution of persistent unobserved heterogeneity can be approximated by a discrete step-wise function, which is sometimes referred to as a discrete factor model (DFM) (Heckman and Singer, 1984; Mroz, 1999). Thus, the estimation procedure seeks to determine (i) the number of unobserved types in the population, K; (ii) the share of the population that is described by each type, θ_k for $k=1,...,K$ where $\sum_{k=1}^{K}\theta_k=1$; and (iii) the impact that each unobserved type k has on all model choices and outcomes, $\mu^k=\{\mu_1^k,...,\mu_{12}^k\}\forall k=1,...,K$.

Conditional on knowing the number of unobserved types in the population, K, the estimation of $\{\theta_0^k,\mu_0^k\}_{k=1}^K$ is straightforward, as these parameters are part of the likelihood function. Determining K is less straightforward. Mroz (1999) recommends an "upwards-testing approach," where one first estimates all model parameters assuming one unobserved type, $K=1$, which produces a log-likelihood function value, LLF_1, and a set of maximizing parameters, $\widehat{\Omega}_1$.[24] The model is then re-estimated with two unobserved types, $K=2$, using the previously estimated parameters, $\widehat{\Omega}_1$, as starting values, which produces a new log-likelihood function value, LLF_2, and a new set of maximizing parameters, $\widehat{\Omega}_2$. A likelihood ratio (LR) test is used to determine whether the additional unobserved type led to a significant improvement in the log-likelihood function. Mroz suggests continuing in this fashion, adding additional points of support so long as significant improvements are made in the value of the log-likelihood function. I amend this rule slightly, requiring both a significant improvement in the log-likelihood function *and* model fit. The model fit criteria is added because the technique is computationally expensive; additional unobserved types increase estimation time substantially. Using this strategy, I arrive at four mass points.

The DFM offers two advantages over a popular alternative, which is to assume a joint parametric distribution (e.g., multivariate normal) over the model's error terms. First, the DFM is more flexible. Mroz (1999), and more recently Guilkey and Lance (2013), uses Monte Carlo simulation in a two-equation, joint MLE setting to show that when the true error distribution is joint normal, DFM estimates are comparable to those derived using the correct distribution. However, when the true error distribution is not normal, the DFM outperforms all other tested estimation methods. Second, the DFM is almost certainly faster than assuming a joint parametric distribution, which typically requires the use of maximum simulated likelihood estimation.[25]

[24] If $K=1$, then $\theta_1=1$ and μ^1 is not separately identified from the model's constants, so the vector is set to zero. Setting $K=1$ assumes that the model's error terms are conditionally independent.

[25] Fox et al. (2011) propose an alternative grid search strategy that guarantees a global maximum and allows for a large number

5.3 Identification

A key challenge to overcome in this research is separately identifying moral hazard from selection effects. Random assignment of health insurance plans is generally considered the *ideal* solution to this problem (Manning et al., 1987). While health insurance is endogenously selected in the MEPS, substantial variation across insurance alternative sets aids in identification in a way that is similar to random assignment. Assuming firm-specific insurance alternative sets are exogenously determined, variation across firms leads observationally equivalent individuals to possess different plans and, therefore, face a different sequence of out-of-pocket medical care prices over the insurance year. In a more reduced-form model of medical care spending, this strategy would be similar to instrumenting for plan generosity with the generosity of the insurance alternative set. In Section 6 of the web appendix, I show how this identification strategy reduces bias in moral hazard estimates in a reduced-form model of medical care demand. Cardon and Hendel (2001) identify moral hazard effects utilizing similar variation in the National Medical Expenditure Survey. Ultimately, while variation in insurance alternative sets aids in identification, this strategy may not fully solve the selection problem, as individuals select their place of employment and, therefore, their insurance alternative set as well.[26]

The dynamic structure of the optimization problem further aids in separately identifying moral hazard effects from selection. Note that the parameters (ω_{00},...,ω_{03}), which measure the marginal utility of consumption, are relatively important in the moral hazard simulation because they dictate price sensitivity. These parameters impact the likelihood function in two places: (i) contemporaneous utility yielded by a medical care choice (Equation 4), which affects medical care choice probabilities (Equation 10) and (ii) the expected future value of insurance (Equation 11), which affects insurance choice probabilities (Equations 13). Were an insurance decision *not* modeled, call this Model A, the parameters determining price sensitivity would be selected only to maximize the probability of observed medical care choices, conditional on prices. However, in a model that includes an insurance decision, call this Model B, price sensitivity parameters are selected to maximize the probability of both observed medical care *and* insurance choices. If health insurance in the population is largely adversely selected, then the insurance choice in Model B puts downward pressure on the price sensitivity parameters from Model A. In other words, were Model B simulated with Model A parameters, the model would over-

of types. Unfortunately, the large parameter set in this model makes implementing this procedure computationally infeasible.

[26] Cardon and Hendel (2001) show in their setting that neither observable demographics nor health explain whether an individual's employer offers *any* health insurance, conditional on income, industry, and occupation, which provides some evidence that employer provider health insurance plans are exogenously determined.

predict insurance take-up. The opposite is true in a world where insurance is largely advantageously selected.

There are two threats to this identification strategy. First, the moral hazard estimate is sensitive to functional form assumptions, which is a common criticism of structural modeling. Ideally, one would estimate the model under a variety of assumptions to verify that results are robust; however, the complexity of the DPP makes estimating the model computationally intensive, preventing this exercise. Thus, structural research typically relies more heavily on economic theory. Second, it is assumed that the unexplained determinants of insurance decisions, ϕ^j, are orthogonal to the unexplained determinants of future illness, medical care prices, and medical care demand. To evaluate this assumption, consider the determinants of insurance decisions that are *explained*. Observed demographics, estimated unobserved types, and observed illnesses at the time of an insurance decision, which the model suggests are predictive of future illnesses, prices, and medical preferences, all impact the expected future value of a plan, as they enter the initial state vector $\mathbf{S_0}$ in Equation 12, meaning they are not contained in ϕ^j.[27] All other determinants are contained in ϕ^j and the model assumes that ϕ^j is uncorrelated with future shocks. There are conceivable threats to this assumption and results should be interpreted with these in mind. For example, a foreseeable, large, isolated medical expense, such as a relatively safe surgery requiring little post-op care, is likely to (i) influence one's insurance decision and (ii) lead to a one-month, unexplained increase in hospital prices, but is unlikely to be captured by the permanent unobserved types.

6 Results

6.1 Parameter Estimates

Table 2 reports estimated preference parameters. A constant relative risk aversion (CRRA) parameter, RA, can be calculated for each individual using $\{\omega_{00},\omega_{01},\omega_{02},\omega_{03}\}$.[28] Risk aversion is found to be decreasing

[27] For example, my findings suggests that relative to women, men are less risk averse, have lower preferences for any medical care, and are less likely to become ill; thus, the model would predict that men are more likely to select low cost, less generous plans. Similarly, the model would predict that unobserved type 1 individuals, who have the strongest preferences for medical care, are likely to select more generous plans. Note the value, then, of capturing permanent unobserved heterogeneity over the insurance year, as it relates to separating moral hazard effects from selection. By allowing for variation in unobserved types, the model can capture heterogeneity in medical care preferences and allow an individual's personal knowledge of their preferences to impact their valuation of an insurance plan.

[28] The CRRA parameter is calculated as $RA = [1 - \omega_{00} - \omega_{01} * age - \omega_{02} * nonwhite - \omega_{03} * male]$. Age is scaled so that the youngest individual (19) has an age of 0. A 40 year-old, then, has an age of 21.

Table 2: Preference Parameter Estimates

	Parameter	Estimate	S.E.
Utility Function			
RA constant	ω_{00}	0.0594	0.0066
RA age	ω_{01}	0.0010	0.0004
RA non-white (black or Hispanic)	ω_{02}	0.0040	0.0114
RA male	ω_{03}	0.0213	0.0139
acute illnesses†	ω_{10}	-28.5934	4.4581
acute illnesses (squared)	ω_{11}	2.7975	0.7919
acute illnesses*age	ω_{12}	0.1894	0.0370
chronic illnesses†	ω_{20}	-3.5427	5.208
chronic illnesses (squared)	ω_{21}	-0.7278	0.6208
chronic illnesses*age	ω_{22}	-0.1531	0.1761
doctor visits	ω_{30}	-1.5748	0.1832
doctor visits (squared)	ω_{31}	-0.1772	0.0276
doctor visits*age	ω_{32}	-0.0008	0.0012
doctor visits*male	ω_{33}	0.0408	0.0203
hospital days	ω_{40}	-5.4660	0.4032
hospital days (squared)	ω_{41}	0.8503	0.0986
hospital days*age	ω_{42}	0.0082	0.0034
hospital days*male	ω_{43}	0.1998	0.0586
any Rx consumption	ω_{50}	-5.8783	0.5979
any Rx consumption*age	ω_{51}	0.0258	0.0068
any Rx consumption*male	ω_{52}	-0.3121	0.1361
any consumption	ω_{60}	-2.1247	0.1151
any consumption*male	ω_{61}	-0.4811	0.0753
negative consumption‡	ω_{70}	0.1000	
Other			
discount factor‡	β	0.9960	
log-likelihood value§	$L(\Omega)$	-56,885.2	

† Parameter scaled by factor of 10 in estimation to improve numerical optimization.
‡ These parameters are not estimated. Note that β is set to 0.996, instead of the traditional 0.95, because decisions are made monthly in the model. $0.996^{12} \approx 0.95$. The parameter ω_{70} measures the (linear) disutility of each dollar of negative consumption (i.e., outspending monthly income, requiring an individual to use savings or to borrow) and is not estimated due to weak identification. Other model parameters are not sensitive to different fixed values for this parameter.

in age, while non-whites and males are found to be less risk averse than whites and females. At the sample mean, $RA = 0.91$, which is just below the Blau and Gilleskie (2008) estimate of 0.96. Cohen and Einav (2007) and Handel (2013) also estimate risk preferences in an insurance contract setting, but assume that individual preferences are constant in the level of absolute risk aversion. To compare our estimates, I calculate the x that makes an individual indifferent to a 50/50 gamble where he either wins \$100 or looses \$$x$ (i.e., the x that solves $u(w) = \frac{1}{2}u(w+100) + \frac{1}{2}u(w-x)$). I estimate $x = \$96.3$ at the sample median, while Cohen and Einav (2007) and Handel (2013) estimate $x = \$99.7$ and $x = \$96.3$, respectfully.

Parameters $(\omega_{10}, \ldots, \omega_{22})$ capture *contemporaneous* disutility from acute and chronic illness. The

parameters suggest that for a 40-year-old individual, a single acute illness causes roughly three times as much contemporaneous disutility as a single chronic illness; however, the expected lifetime disutility due to a single chronic illness is larger than that of an acute illness by several orders of magnitude.[29] This finding is consistent with the prior that most acute illnesses cause intense short-term discomfort (e.g., influenza, bronchitis) but have little impact on future health or utility, while many chronic illnesses cause less short-term discomfort (e.g., hypertension, heart disease) but shorten one's expected lifespan, dramatically reducing expected lifetime utility. The estimates also suggest that the disutility of acute illness is decreasing in age and in the number of acute illnesses, while the disutility of chronic illness is increasing in age and the number of chronic illnesses.

Parameters (ω_{30},...,ω_{61}) capture the *net* direct effect of medical care consumption on utility - the physical, psychological, and time cost of medical care consumption may have negative effects on these parameters, while some individuals may enjoy consuming medical care, independent of its productive health effects, which has a positive effect on these parameters. Interpreting the linear and quadratic consumption parameters is not useful without also considering the unobserved heterogeneity parameters discussed below. However, these parameters do reveal that, relative to women, men have significantly lower preferences for prescription drugs and any medical care consumption, yet stronger preferences for doctor and hospital care. The parameters also suggest that preferences for hospital visits and prescription drugs are increasing in age.

The parameters in Table 3 describe the discrete step-wise function used to approximate the joint distribution of unobservables in the model. Recall that the technique identifies the effect that each unobserved type has on the model's choices and outcomes, along with the probability of being a particular type. For identification, I fix the effect and probability parameters of Type 1 individuals to zero. Estimation reveals that 51 percent of the population is Type 1. The medical care preference parameters in Table 2 are fully representative for these individuals, meaning preferences for doctor, hospital, and prescription drug consumption are negative. Individuals represented by the other three types are sicker, have greater distaste for doctor and hospital visits, and receive lower price draws for doctor and hospital visits (on average) than Type 1 individuals. Type 4 individuals are unique in their distaste for prescription drugs and abnormally low doctor's office price draws.[30]

[29]The relative size of ω_{10} and ω_{20} can be misleading. Note that the disutility of acute (chronic) illness is decreasing (incresing) in age; ω_{10} and ω_{20} would be closer in size without the age interaction. Moreover, using parameters (ω_{10},...,ω_{22}) and the closing function parameters (Table A25 of the web appendix), one can determine that the expected lifetime disutility from a single chronic illness is almost nine times greater than that of a single acute illness.

[30]The impact that unobserved heterogeneity has on parameter estimates, model fit, and moral hazard effects is discussed

Table 3: Permanent Unobserved Heterogeneity Parameter Estimates

	Param.	Type 1	Type 2		Type 3		Type 4	
			Est.	S.E.	Est.	S.E.	Est.	S.E.
Mass Point Location								
acute illness prob.	μ_1	0.0	0.569	0.059	0.556	0.066	0.573	0.064
chronic illness prob.	μ_2	0.0	0.985	0.099	0.978	0.101	1.109	0.106
doctor visit pref. (linear)	μ_{3a}	0.0	-1.599	0.144	-1.636	0.150	-1.336	0.169
doctor visit pref. (squared)	μ_{3b}	0.0	0.423	0.025	0.397	0.028	0.412	0.025
hospital day pref. (linear)	μ_{4a}	0.0	-1.536	0.274	-1.543	0.297	-1.352	0.323
hospital day pref. (squared)	μ_{4b}	0.0	0.243	0.082	0.230	0.097	0.282	0.086
any Rx pref.	μ_5	0.0	0.788	0.549	0.759	0.563	-3.788	0.652
any consumption pref.	μ_6	0.0	1.213	0.227	0.832	0.275	1.631	0.146
doctor visit price dist.	μ_7	0.0	-3.161	2.402	-2.816	2.729	-13.005	2.330
hospital day price dist.	μ_8	0.0	-0.120	0.198	-0.110	0.213	-0.445	0.180
monthly Rx price dist.	μ_9	0.0	0.770	0.032	-0.747	0.033	-0.090	0.040
initial acute illness prob.	μ_{10}	0.0	1.360	0.195	1.155	0.234	0.899	0.232
initial chronic illness prob.	μ_{11}	0.0	2.300	0.222	2.102	0.254	1.460	0.251
initial health status prob.	μ_{12}	0.0	0.387	0.173	0.385	0.208	0.192	0.212
Type Probabilities†								
parameter estimate	θ	0.0	-0.961	0.094	-1.363	0.106	-1.129	0.122
type probability		51.0	19.5		13.0		16.5	

† Probabilities are derived from estimating parameter θ^k such that $Pr(\mu^k) = \frac{exp(\theta^k)}{\sum_{k'=1}^{4} exp(\theta^{k'})}$.

Estimating the productive effects of medical care on health and/or illness is challenging for a number of reasons. A simple regression of wellness on medical care consumption usually reveals negative correlation, which likely reflects bias associated with selection into consumption (i.e., only sick people consume care) and omitted controls for lagged wellness. In this research, I address the first issue by modeling medical care consumption and allowing for common unobserved individual heterogeneity that affects both medical care decisions and health outcomes. I address the second issue by simultaneously controlling for the number of acute illnesses and chronic illnesses entering the month, as well as general health status at the beginning of the insurance year. Table 4 reports acute and chronic illness probability parameter estimates. I find evidence of medical care productivity across all types of care and illness. Both acute and chronic illness probabilities decrease with additional doctor and hospital visits (α_{40},α_{42},δ_{40},δ_{42}); however, each additional visit is less productive than the previous (α_{41},α_{43},δ_{41},δ_{43}). Prescription drugs seem to be the most productive form of medical care in curing and preventing illness (α_{44},δ_{44}), but these parameters are difficult to interpret given that prescription drugs are only consumed on the extensive margin.[31]

at length in Section 11 of the web appendix.

[31] Price, initial condition, and closing function parameters are presented in Section 8 of the web appendix.

Table 4: Illness Probability Parameter Estimates

	Acute			Chronic		
	Param.	Estimate	S.E.	Param.	Estimate	S.E.
constant	α_{00}	-2.1088	0.0940	δ_{00}	-5.0130	0.1723
male	α_{10}	-0.3296	0.0374	δ_{10}	-0.0182	0.0281
non-white (black or Hispanic)	α_{11}	-0.0824	0.0373	δ_{11}	-0.0173	0.0130
education (highest grade completed)	α_{12}	0.0016	0.0053	δ_{12}	0.0188	0.0044
age	α_{13}	-0.0030	0.0018	δ_{13}	0.0047	0.0044
lives in a MSA	α_{14}	-0.0016	0.0465	δ_{14}	-0.0326	0.0177
income (in 1996 dollars)	α_{15}	0.0045	0.0011	δ_{15}	0.0002	0.0004
initial health	α_{16}	0.0081	0.0156	δ_{16}	0.0546	0.0114
acute illness	α_{20}	3.2414	0.0560	δ_{20}	-0.2161	0.0445
acute illness (squared)	α_{21}	-0.1965	0.0160	δ_{21}	0.0561	0.0112
consecutive mons. with acute illness†	α_{22}	0.0179	0.0131	δ_{22}	-0.0143	0.0067
consecutive mons.* initial month	α_{23}	0.0115	0.0122	δ_{23}	-0.0002	0.0044
chronic illness	α_{30}	0.1515	0.0374	δ_{30}	0.3153	0.0444
chronic illness (squared)	α_{31}	-0.0201	0.0088	δ_{31}	-0.0137	0.0062
doctor visits	α_{40}	-0.0481	0.0071	δ_{40}	-0.0346	0.0076
doctor visits (squared)	α_{41}	0.0053	0.0009	δ_{41}	0.0028	0.0007
hospital days	α_{42}	-0.0946	0.0181	δ_{42}	-0.0392	0.0129
hospital days (squared)	α_{43}	0.0193	0.0040	δ_{43}	0.0072	0.0031
Rx consumption	α_{44}	-0.1475	0.0212	δ_{44}	-0.1869	0.0378
acute illness*doctor visits	α_{50}	-0.0058	0.0010			
acute illness*hospital days	α_{51}	-0.0016	0.0021			
acute illness*Rx consumption	α_{52}	-0.0041	0.0040			
chronic illness*doctor visits	α_{53}	0.0013	0.0006			
chronic illness*hospital days	α_{54}	-0.0017	0.0014			
chronic illness*Rx consumption	α_{55}	0.0063	0.0036			
threshold 2	κ_2^a	3.2557	0.0415	κ_2^c	2.8027	0.2648
threshold 3	κ_3^a	6.0154	0.0727			
threshold 4	κ_4^a	8.2134	0.1261			

Notes: Month indicators are included as regressions but are not reported here. Medical care consumption types are not interacted with lagged illness states in the additional chronic illness model because there is not enough variation in the data to identify the parameters.

† I control for the number of consecutive months with *any* acute illness leading up to the current month to capture the impact of a more severe acute illnesses on illness transition probabilities. Due to the nature of the data, this count begins in the first month of the insurance year, meaning an individual entering the third month of the year with two consecutive acute illness months may actually have had an acute illness even longer. To control for this measurement error, I interact consecutive months with an acute illness with an indicator that equals one if an individual has had an acute illness in every month since the first month of the year.

In summary, the estimated parameters support a set of theoretical priors used to develop this economic model of dynamic decision making. Specifically, the parameters suggest that the contemporaneous effect of medical care consumption on utility is negative on average - i.e., individuals have negative preferences for medical care and the consumption of medical care reduces one's ability to consume non-medical goods. Individuals are found to consume medical care because it promotes wellness in future months, from which they derive utility, and because of heterogeneous unobserved preference shocks. Note that a formal analysis of model fit is presented in Section 9 of the web appendix. In this section, I argue that the model can explain unique features

of the data, which suggests that the model adequately represents the true unknown data generating process.

6.2 Moral Hazard

I define moral hazard as the percentage of mean annual medical expenditure that would not occur in the absence of insurance. To calculate this statistic, I use the estimated model to forward simulate the behavior of individuals when they are uninsured (i.e., are responsible for the full costs of medical care consumed). Next, I simulate behavior when individuals are required to select a health insurance plan from the alternative set offered by their employer.[32] I consider two scenarios when insured: either an individual selects his most preferred plan or he is forced into *full* insurance coverage (i.e., he pays zero dollars out of pocket at all times during the insurance year.) Mean expenditure is \$663 when uninsured and \$1,435 when insured in the preferred plan. Column 1 of Table 5 reports the distribution of the difference in total annual medical expenditure when insured by the preferred plan versus uninsured. The expenditure induced by moral hazard amounts to 53.1 percent (s.d. = 4.0), on average, of total annual medical expenditure when insured. Put differently, individuals consume 2.16 times more medical care dollars when insured than when uninsured. These moral hazard effects vary widely across the population. Upon winsorizing the insurance-induced spending distribution at the 99th percentile (i.e., spending increases are capped at \$10,906), the percentage of mean expenditure explained by moral hazard falls to 46 percent (s.d. = 3.1). Furthermore, 54.4 percent (s.d. = 10.3) of the sample does not increase their spending at all in response to coverage. Column 2 reports the predicted increase in expenditure when individuals are moved from no coverage to full coverage.

The moral hazard estimate relates directly to how well adverse selection is measured in the model. For example, as is discussed in Section 5.3, modeling unobserved heterogeneity aids in capturing adverse selection effects, as permanent unobservables such as strong individual preferences for medical care or private knowledge of an existing/impending illness are positively correlated with both generous insurance selection and medical care spending. In fact, as I move from 1 to 2 to 4 unobserved types in estimation, the left tail of the insurance-induced spending distribution thickens considerably, while the moral hazard effect at the mean falls from 57.4% (s.d. 3.1) to 56.1% (s.d. 4.2) to 53.1% (s.d. 4.0). As I argue in Section 11.2 of the web appendix, these findings suggest that the unobserved types identify sub-populations that are unlikely to use medical care, even when ill (i.e., types 2 and 3, but type 4 in particular). As such, these individuals are less

[32] In both stages, for each individual I sample 100 times from the joint error distribution, permanent unobserved heterogeneity distribution, and multivariate normal parameter distribution.

responsive to changes in the out-of-pocket cost of medical care that is associated with insurance acquisition.

Table 5: Predicted Increase in Annual Expenditure

	WYDM Model		RAE Model	
Percentile	Force	Full	Force	Full
50	0	0	0	0
55	17	69	0	0
60	83	144	0	84
65	164	250	48	203
70	285	403	176	380
75	462	616	346	630
80	722	927	602	987
85	1,143	1,414	1,004	1,529
90	1,904	2,265	1,720	2,467
95	3,651	4,172	3,437	4,791
99	10,906	12,044	12,295	16,266
Mean	772	899	760	1,115

Notes: This table displays percentiles of the distribution of the predicted increase in mean annual medical care that results from insurance acquisition. In the first and third columns, labeled *Force*, an individual is forced into his most preferred plan among those offered by his employer. In the second and fourth columns, labeled *Full*, an individual is forced into full coverage, with an out-of-pocket insurance premium of zero. These expenditure values are then compared to those calculated in a third counterfactual, where individuals are forced to be uninsured.

Columns 1-3 of Table 6 report changes in medical care consumption, prices, and illness as individuals move from uninsured to holding their preferred plan. Insurance acquisition increases the consumption of all three types of medical care. The most dramatic percentage increase is in total annual hospital consumption, though the level increase is small (i.e., an additional 0.67 hospital visits a year on average). The increase in prescription drug consumption is small because drugs are both inexpensive and productive, making them the most attractive form of medical care consumption for the uninsured.[33] Insurance acquisition also increases the average price paid for both doctor and hospital visits, despite the fact that an individual cannot observe prices. This finding suggests that an individual who expects higher price draws consumes care when insured that he would not when uninsured, ceteris paribus. I also find that the increase in medical care consumption that results from insurance acquisition has only a small impact on illness. The average sum of acute illness months (i.e., $\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{12}A_{it}$) and average number of chronic illnesses at the end of the year fall by only 1.16 and 0.23 percent, respectively.

Several alternative measures of moral hazard effects are reported in the literature. By simulating behavior

[33]In Section 10 of the web appendix, I compare my findings on price sensitivity across types of medical care to those found in the literature.

Table 6: Change in Endogenous Variables with Insurance Acquisition

Variable	WYDM Model Uninsured	WYDM Model Insured	WYDM Model % Δ	RAE Model Uninsured	RAE Model Insured	RAE Model % Δ
Medical Care Consumption						
average annual doctor visits	3.85	5.41	+40.52	4.17	4.97	+19.18
average annual hospital visits	0.05	0.72	+1,340.00	0.33	0.72	+118.18
average annual Rx months	4.15	4.27	+2.89	4.34	4.46	+2.76
Medical Care Prices						
average doctor price	88.56	93.95	+6.07	75.85	90.57	+19.41
average hospital price	610.94	869.03	+42.24	267.54	869.69	+225.07
average Rx price	61.01	60.52	-0.80	65.65	68.92	+4.98
Illness†						
total acute illness months	7.75	7.66	-1.16	*	*	*
average number of chronic illnesses by years end	0.860	0.858	-0.23	*	*	*

† Illness transitions are treated as exogenous in the RAE model, so there is no change due to insurance acquisition.

under various insurance conditions, I am able to calculate these measures using the estimated WYDM model and MEPS data. Comparisons are reported in Table 7. Row 1 contains the estimates of Manning et al. (1987) who use the experimental RAND HIE data to estimate a co-insurance (arc) elasticity of medical care demand of 0.17 for the 0-25 percent co-insurance range and 0.22 for the 25-95 percent co-insurance range. To calculate comparable measures, I simulate the within-year decision-making model under three cost-sharing arrangements. All three arrangements feature no deductible, no MOX level, and no premium; differing only by a universal co-insurance rate, which is set to 0, 25, or 95 percent.[34] I estimate a co-insurance elasticity of medical care demand of 0.20 for the 0-25 co-insurance range and 0.35 for the 25-95 co-insurance range. The remaining three comparisons are described in the footnotes of Table 7.[35]

While the above exercise highlights the flexibility of the WYDM model to produce a variety of moral hazard estimates, the measures reported in Table 7 are not truly comparable, as each study uses a different sample population in its empirical analysis. These populations differ by observable characteristics, likely differ by unobservable characteristics, and allow health insurance to be selected from different sets of alternatives;

[34]The RAND HIE plans used for price elasticity estimation had no premium or deductible. The plans did feature a MOX level, so researchers estimate demand by individuals more than $400 from their MOX level to avoid the price distortion (Manning et al., 1987). Rather than recreating this approximation technique, I eliminate the MOX level in simulation.

[35]I estimate the increase in mean annual expenditure caused by switching from no insurance to full coverage to be 135 percent, which is large compared to Einav et al. (2013). However, note that the same simulation results can be used to calculate the percentage of full insurance expenditure not explained by moral hazard (i.e., 663/[663+899] = 42 percent), which is near the preferred measure of Keeler and Rolph (1988), who study the RAND HIE data.

Table 7: Other Measures of the Effects of Moral Hazard

	Measure	Reported Estimate	Within-Year Model Estimate
Manning et al. (1987)[a]	Co-insurance arc elasticity of medical care demand 0%→25%, 25%→95%[b]	0.17, 0.22	0.20, 0.35
Keeler and Rolph (1988)	Percentage of full insurance expenditure not explained by moral hazard[c]	55%	42%
Einav et al. (2013)	Percentage increase in mean annual expenditure, no insurance to full coverage[d]	30%	135%
Bajari et al. (2014)	Percentage of preferred insurance expenditure explained by moral hazard[e]	45%	24%

a These results are also reported in Keeler and Rolph (1988).
b The arc elasticity is calculated as $E_{arc}=((q_2-q_1)/(p_l-p_1))\times((p_2+p_1)/2)/((q_2+q_1)/2)$, where q is mean annual medical expenditure and p is the co-insurance rate. Manning et al. (1987) make this calculation for each type of care and then weight elasticities for various types of care by share of spending.
c q_2/q_1 is calculated, where q_2 is mean annual medical expenditure with a 95% co-insurance rate (i.e., near no insurance) and q_1 is mean annual medical expenditure with full insurance.
d $(q_2-q_1)/q_1$ is calculated, where q_2 is mean total annual medical expenditure for the population under full coverage and q_1 is mean total annual medical expenditure for the population under no coverage.
e $(q_{2i}-q_{1i})/q_{2i}$ is calculated for each individual, where q_{2i} is total annual medical expenditure with preferred/chosen coverage and q_{1i} is total annual medical expenditure with no coverage for individual i. Note that this measure differs from my preferred estimate, which compares mean expenditure levels, rather than the average of individual spending ratios.

all of which are likely to impact the estimated effect of moral hazard on medical care consumption.

6.3 A Representative Annual Expenditure Model

The following steps allow me to determine whether an annual medical care decision-making model produces a biased estimate of moral hazard effects. First, I impose a set of assumptions on the WYDM model that are representative of a general annual expenditure model. I refer to this model as the representative annual expenditure (RAE) model. Second, using the same data and imposing the same functional form assumptions as the WYDM model, I estimate the structural parameters of the RAE model.[36] Third, I calculate moral hazard effects via simulation. Fourth, I compare the moral hazard effects produced by the two models.

[36] An alternative strategy would be to specify and estimate a true annual expenditure model, which would require functional form assumptions that differ from the WYDM model, potentially leading to divergent results independent of the economic constraints/incentives imposed on the models.

6.3.1 Model Construction

The RAE model is constructed by imposing three assumptions on the WYDM model. First, I assume that illness transitions are exogenously determined. Each of the annual expenditure models discussed in this paper lack health production, leaving preferences as the only motivation for medical care consumption.[37] Second, I assume that contemporaneous medical care prices are known to an individual. Third, I assume that in each month after the insurance decision is made, an individual knows all of his future price and medical care preference shocks.

Under these assumptions, the RAE model *represents* an annual expenditure model in the sense that an individual solves his optimization problem under the same incentives and information as an annual expenditure model. In an annual expenditure model, an individual makes a health insurance decision without knowing the expenditure shock he receives in the future. In the RAE model, an individual makes a health insurance decision without knowing the medical care consumption and price shocks he receives in the future. In an annual expenditure model, an individual makes an annual medical expenditure decision where (i) the only benefit of medical care is a contemporaneous utility gain and (ii) there is no uncertainty about the current or future costs or benefits of his decision. In the RAE model, the same is true for decisions made over the course of the year.

6.3.2 Empirical Findings

RAE model parameter estimates and model fit analysis can be found in Section 12 of the web appendix. Mean predicted expenditure is \$683 when individuals are uninsured and \$1443 when individuals are covered by their preferred plan; thus, the percentage of mean preferred plan expenditures explained by moral hazard is 51.9 percent (s.d. = 5.9), which is smaller than that produced by the WYDM model (53.1 percent, s.d. = 4.0). Columns 3 and 4 of Table 5 report the distribution of insurance-induced spending. Like the WYDM model, the effect varies widely across the population. For the RAE model, upon winsorizing the distribution at the 99th percentile, the mean moral hazard effect falls to 42.6 percent (s.d. = 5.2), which is significantly smaller than the estimate produced by the WYDM model (46 percent, s.d. = 3.1). Furthermore, 64 percent (s.d. = 8.1) of the sample does not respond to insurance aquisition, which is significantly larger than the group predicted by the WYDM model (54.4 percent, s.d.=10.3).

[37] Khwaja (2010) models endogenous health production; however, medical care decisions are made biennially, so endogenous health transitions do not create dynamic medical care incentives within the health insurance year.

In summary, there are two notable differences in the insurance-induced spending distributions produced by the RAE and WYDM models. First, compared to the WYDM model, the RAE model predicts that more individuals do not increase their medical care consumption at all when they acquire insurance. Provided the discussion in Section 1 of the web appendix, this finding is consistent with a sample population that is risk averse and has a low willingness to pay for medical care. The medical care consumption patterns presented in Table 6 are consistent with this hypothesis. The WYDM model predicts a larger percentage increase in average medical care consumption than the RAE model for all three types of care, with the largest differences for the most expensive types of care (i.e., hospital and doctor visits). The second notable difference is that the distribution of insurance-induced spending produced by the RAE model has a longer right tail. This finding is consistent with Section 1 of the web appendix if a small group of individuals in the sample face the possibility of very large price draws, p_h, yet also place a lot of value on medical care (i.e., are in the right tail of the willingness to pay distribution). According to this theory, when uninsured and facing price uncertainty these individuals are likely to purchase care despite the risk of a high price draw because of their strong preferences for medical care; however, when uninsured and facing known prices these individuals only purchase care when prices are low. Evidence of this behavior can also be see in Table 6, as the WYDM model predicts an increase in the average price paid for doctor and hospital visits, but the increase is much larger in the RAE. In other words, RAE model individuals never purchase expensive care when uninsured, while those in the WYDM model with a large W may purchase care when uninsured and receive a high price draw.

While the two models produce insurance-induced spending distributions that are different statistically, it is not clear whether these differences are relevant economically. As such, I use mean moral hazard estimates from both models to predict the increase in total U.S. medical expenditure that would result from mandating that every uninsured individual in 2013 have health insurance coverage. Average medical care spending among the 40.8 million uninsured individuals in 2013 was \$2,443 (Coughlin et al., 2014). Unadjusted mean moral hazard estimates from the two models suggest increases in average expenditure upon insurance acquisition of \$2,834 (WYDM) and \$2,712 (RAE).[38] With these estimates, the WYDM model predicts an increase in total U.S. expenditure that is about five billion dollars larger than the RAE model, or just over four percent of the WYDM model's predicted expenditure increase, which does not seem like an economically meaningful difference. That said, as was discussed above, the mean moral hazard

[38]For the WYDM model, I multiply \$2,443 by 1.16 [=(1,435/663)-1]; for the RAE model, by 1.11 [=(1,443/683)-1].

estimate is quite sensitive to small changes in the right tail; thus, a more conservative prediction might limit the impact of this right tail. After winsorizing the insurance-induced spending distribution at the 99th percentile, the mean moral hazard estimates generate expenditure increases of $2,711 (WYDM) and $2,442 (RAE). Aggregated up, the WYDM model predicts an increase in total U.S. medical expenditure that is roughly 11 billion dollars larger than the increase predicted by the RAE model, or 10 percent of the WYDM model's predicted expenditure increase, which is a more economically meaningful difference.

6.4 Counterfactual Analysis: An Empirical Test of Grossman 1972

The seminal Grossman (1972) model assumes that the primary motivation for medical care consumption is the production of a durable health capital stock, which produces "healthy days", from which individuals derive utility. Zweifel and Manning (2000) describe an empirical test of Grossman's health capital model, which I conduct using the WYDM model:

> "In the logic of a Grossman 1972 model, an unexpected reduction in price [of medical care] will induce individuals to want to increase their health stock above the preceding optimal level. To bring desired and current health into alignment requires an increase in medical care demand ... Once the new desired stock equals the current, demand would fall back to a lower level."

Specifically, I simulate insurance acquisition for each individual in my sample and examine the change in illness and spending patterns over the insurance year.[39] My findings are presented in Figures 2a and 2b. Both figures display on the horizontal axis - the month of the insurance year - and on the vertical axis - the average difference between fully insured and uninsured individuals for an outcome of interest. Figure 2a shows that in every month of the year, average medical care consumption is higher for insured individuals. Figure 2b shows that this medical care consumption leads the average individual to possess fewer illnesses when insured than when uninsured. Moreover, over the course of the year, the health benefit of this additional medical care consumption accumulates. In each consecutive month, the average number of acute and chronic illnesses possessed by insured individuals decreases relative to uninsured individuals. In Figure 2a, one can see that for doctor's office visits, the additional consumption due to insurance possession

[39] Note that the WYDM model is sufficiently flexible to reject the health capital hypothesis, despite being motivated by Grossman's model. Medical care consumption enters the utility function directly, as well as affecting the illness transition probabilities. As such, whether the primary motivation for medical care consumption is a desire to improve health or is preference related is an empirical question.

falls over the course of the year as health improves. This pattern is not observed for hospital or prescription drug care.[40] In other words, the WYDM model predicts the exact patterns of illness and doctor's office consumption that would be predicted by the Grossman (1972) health capital model.

6.5 Counterfactual Analysis: Insurance Alternative Set Generosity

As medical expenditure per capita continues to grow in the U.S., many employers have sought to limit their financial exposure by offering a less generous basket of health insurance alternatives. According to a 2016 report by Towers Watson, over three quarters of employers currently offer a high deductible health insurance plans (HDHP).[41] Moreover, from 2010 to 2016 the percentage of firms *only* offering HDHPs grew from 4 to 20 percent, with predicted growth to 24 percent by 2017. The growing popularity of HDHPs should have different welfare implications across income and health distributions, which can be examined using the WYDM model.

In this section, I analyze individual behavior under two counterfactual insurance settings. The first counterfactual adds a HDHP to every individual's alternative set. Abstracting from supply-side effects, one should expect the inclusion to benefit uninsured individuals who have been priced out of the market and overinsured individuals who are currently constrained by their employers' limited offerings. The medical care spending effects of this experiment are ex-ante ambiguous, as one would expect the first group to spend more and the second group to spend less. The second counterfactual replaces the most generous current health insurance offering with a HDHP. Compared to the first counterfactual, one should expect a reduction in both spending and consumer welfare, as those previously holding generous coverage are forced into less-generous plans; however, the net spending and welfare effects are again left as an empirical question.

The HDHP used in this analysis has an annual deductible, doctor and hospital co-insurance rate, and annual MOX level of \$582, 20%, 24%, and \$3,287 (in 1996 dollars), respectively, which represents the 95th percentile of these characteristics for plans observed in the data.[42] The total premium is imputed

[40] Regarding to prescription drugs, I also find that consumption is not very sensitive to changes in prices (see Table 6), so small differentials due to small changes in illness would be difficult to detect. Moreover, my coarse measure of health capital (i.e., number of acute and chronic illnesses) may miss important heterogeneity in illness severity and accidents, which may be less responsive to early-year increases in medical care consumption. As such, one might expect hospital care to be less responsive to changes in illness counts than other types of care.

[41] In 2016, the IRS defined a HDHP to have a minimum (single) deductible of \$1,300 and an MOX level of \$6,550.

[42] An earlier version of this paper defined a HDHP as the average bronze plan offered in 2016 in ACA marketplaces. The plan's deductible was larger than 99% of the deductibles observed in the estimation sample, raising concerns about

using a simple linear model that projects total annual premiums observed in the data onto a set of plan characteristics and interactions. The exercise produces an annual premium of $1,679, which is just under the 20th percentile of premiums observed in the data. To determine the out-of-pocket premium that each individual pays for the HDHP, I calculate the maximum subsidy that their employer offers among its existing plans, and assume the HDHP would be subsidized by this amount. As a result, just 45 percent of the sample faces a non-zero out-of-pocket premium for the HDHP.

The first two columns of Table 8 describe the switching behavior of the previously uninsured population once the HDHP is made available. In general, uninsured individuals are relatively low earning and receive a health insurance offer with few alternatives and higher than average minimum premiums. This characterization is especially true for the subsample that elects to switch into the HDHP, suggesting that these individuals were previously priced out of the market. The next two columns of Table 8 describe the switching behavior of the previously insured population. Compared to non-switchers, switchers appear to be overinsured - they have better initial health, lower income, and are offered fewer alternatives to choose from. Presumably, the HDHP offering is attractive to both column 2 and column 4 individuals because they currently are not offered a low-cost high-risk plan, which suits their preferences.[43] Total expenditure increases (decreases) by an average of 19 percent for previously uninsured (insured) individuals purchasing the HDHP. Since there are nearly five and a half times more insured (than uninsured) switchers, the net effect for switchers is a nearly 17 percent reduction in total expenditures. Overall, the policy leads to an average reduction in total annual medical expenditure in the population of roughly $37 per individual, or 2.9 percent. Setting aside supply-side effects, adding the HDHP has positive welfare effects for all switchers. The average switcher (24 percent of the population) would need to be paid $194 to be just as well off prior to the addition of the plan as they are after the HDHP is offered.

The second counterfactual replaces every individual's most generous health insurance offering with the HDHP described above. I identify the most generous current offering as the plan requiring the lowest out-of-pocket expenditure given $2,000 in total medical expenditure.[44] Results are summarized in columns

out-of-sample fit. I thank an anonymous referee for their insights on this experiment.

[43]To understand exactly how much less coverage the HDHP provides individuals switching from an observed plan to the HDHP, I simulate $2,000 in total medical expenditures under both plans. Out-of-pocket expenditures are $324 under the observed plan and $894 under the HDHP, so roughly 2.7 times higher. That said, for previously insured switchers, the HDHP's total annual premium averages $425 less than their previously selected plan.

[44]I split expenditure equally between doctor and hospital visits. Results are robust to expenditure levels of $1,500 and $5,000.

5 and 6 of Table 8, which examines the policy's effect by insurance status after the first counterfactual (i.e., after the HDHP has been added). Compared to other insured individuals (column 5), those who are forced out of the most generous coverage (column 6) enter the year with a greater number of illnesses, have less income, and are more likely to be of unobserved types 2-4, which are associated with the worst health outcomes. Those forced out of the most generous coverage also select from a smaller set of plans, which is somewhat mechanical; however, the table shows that generous plans require nearly half as much out-of-pocket spending for $2,000 of medical expenditure as non-generous plans - $251.54 vs. $536.43, respectively. Total annual medical expenditure decreases by an average of $271.89 (19 percent) for individuals forced out of generous coverage, a finding generally consistent with that of Brot-Goldberg et al. (2017).[45] As expected, this policy is welfare reducing for those wishing to purchase generous coverage (about 38 percent of the population). These individuals could lose $271.88 prior to the removal of the generous plan and be just as well off as they are after the plan is removed. In total, the welfare decline associated with removing the most generous plan is roughly 2.2 times larger than the welfare gain associated with adding the HDHP.

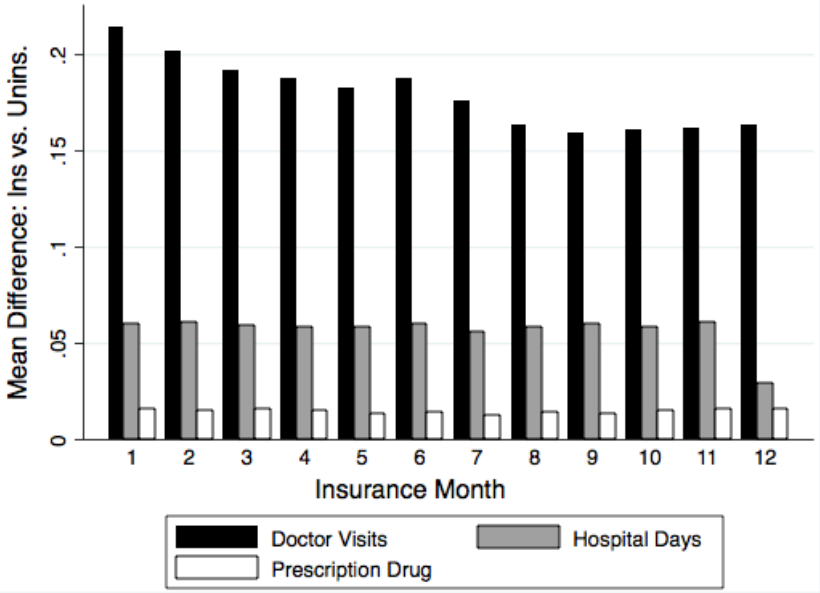
In summary, the above counterfactuals highlight the trade-off that employers face in transitioning to less generous health insurance alternative sets. The policy change has the potential to reduce total annual medical expenditure and improve welfare for some, while substantially reducing welfare for others. In this particular setting, the (consumer) welfare losses significantly outweigh the gains. That said, it is important to interpreted these results with the limitations of the model in mind. The model does not internalize the effects that changes in insurance and medical care demand have on risk pools and, therefore, insurance premiums, which is the topic of Cutler and Reber (1998).[46] The model also predicts behavior for a single year, which ignores any concern about the long-run health implications of less-generous insurance alternative sets.

[45] Brot-Goldberg et al. (2017) estimate a 11.8 - 13.8 percent reduction in total medical care spending when employees were moved from a zero cost-sharing plan into a HDHP covering an average of 76 percent of total annual expenditures. In my experiment, individuals with generous coverage were forced out of a plan covering 76 percent of expenditure on average into a plan covering 50 percent, so while the coverage levels differ between the two experiments, the change in coverage is similar. The fact that the Brot-Goldberg et al. (2017) sample is comprised of high earners, who may be less price sensitive than the average individual, could explain my larger estimates.

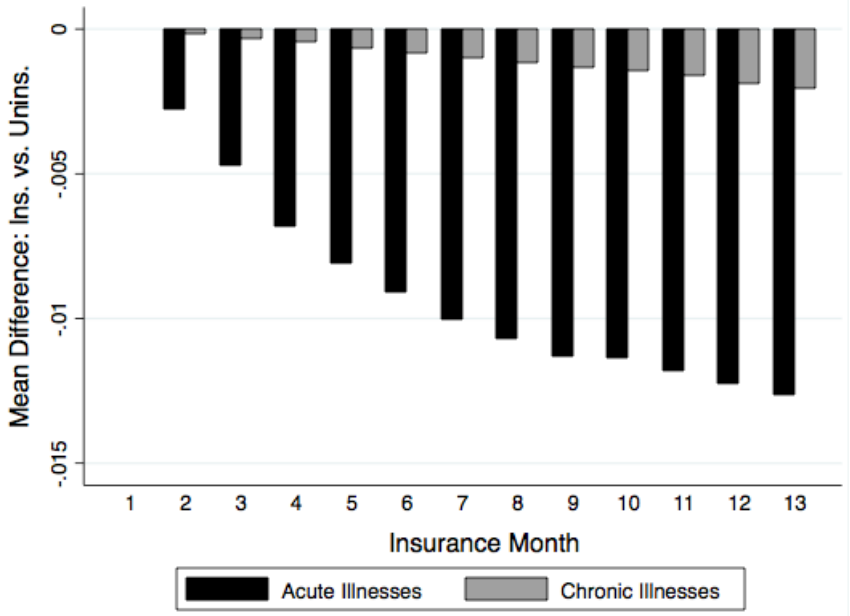
[46] These general equilibrium effects would likely increase the estimated welfare loss. In the first counterfactual, insurance premiums of existing plans are likely to rise as relatively healthy individuals move into the HDHP, reducing welfare for individuals remaining in the old plans. In the second counterfactual, insurance premiums of all plans are likely to rise as the risk pool receives an influx of unhealthy individuals who previously purchased generous coverage.

Figure 2: Medical Care and Illness Differences by Insurance Month

(a) Medical Care Consumption



(b) Illness



Notes: These figures compare the difference in (a) mean medical care consumption and (b) mean number of illnesses between fully insured and uninsured individuals in each month of the insurance year. In both figures, the horizontal axis measures the number of months since the beginning of the health insurance year. In Figure (a), the vertical axis measures, for example, the average number of doctor's office visits for a fully insured individual minus the average number of doctor's office visits for an uninsured individual in each month. For prescription drugs, the vertical axis measures the difference in the proportion of insured and uninsured individuals consuming any drugs. The figure illustrates that the increase in doctor's office visits that results from insurance acquisition decreases over the course of the year, presumably due to an improvement in health. In Figure (b), the vertical axis measures the average number of illnesses for a fully insured individual minus the average number of illnesses for an uninsured individual in each month. The figure illustrates that the full health improvement caused by insurance acquisition is not immediate, but accumulates over time. Note that both images are produced from simulated data, so that the insured vs. uninsured comparison is done for the same set of individuals.

Table 8: Characteristics of Insurance Switchers vs. Non-Switchers

	Counterfactual 1				Counterfactual 2	
	Uninsured		Insured		Insured	
	No Switch	Switch to HDHP	No Switch	Switch to HDHP	Non-Generous	Generous
Δ in welfare (CV)	*	175.60	*	196.81	*	-271.88
Δ in total expenditure	*	57.21	*	-194.35	*	-271.89
Δ in OOP expenditure	*	-100.99	*	167.31	*	181.39
Δ in doctor visits	*	0.61	*	-0.19	*	-0.40
Δ in hospital visits	*	0.11	*	-0.20	*	-0.28
Δ in Rx months	*	0.14	*	-0.02	*	-0.03
Characteristics of offered alt. set						
Number of plans †	1.94	1.43	5.29	2.09	6.8	3.14
Min. OOP premium	430.80	509.64	250.44	299.16	287.64	272.88
Total premium, selected	0	0	2,040.73	2,103.94	1,921.48	1,989.80
OOP premium, selected	0	0	336.02	341.74	308.40	306.07
OOP payment for $2,000 expenditure, selected	2,000	2,000	261.03	324.54	536.43	251.54
Acute illnesses entering year	52.1	50.4	60.3	53.7	57.3	60.2
Chronic illnesses entering year	48.4	46.8	66.3	36.8	53.1	67.3
Income	2,336.16	2,186.13	2,916.72	2697.24	2,906.06	2745.81
Unobserved type						
Type 2:	12.8	13.2	20.6	18.0	18.7	21.4
Type 3:	11.9	10.3	13.7	11.4	12.4	14.1
Type 4:	10.9	11.3	18.4	13.5	15.3	19.6
Observations	7,880	4,463	86,101	24,756	68,737	46,583
Percentage of Population	63.8	36.2	77.7	22.3	59.6	30.4

† In all columns, number of plans refers to the number of plans actually offered by the employer.

* Notes: In the first two columns, I compare the characteristics of uninsured individuals who remain uninsured to uninsured individuals who decide to purchase the HDHP when offered. In the second two columns, I compare insured individuals who remain in their plan to insured individuals who switch to the HDHP when offered. In the final two columns, I limit the sample to all individuals selecting any plan in Counterfactual 1 (i.e., when the HDHP is available). I then compare those selecting the most generous offered coverage to those selecting a plan other than the most generous one they are offered.

7 Conclusion

A recent literature focused on optimal health insurance plan (and alternative set) design has utilized structural models of insurance decisions and medical expenditure (e.g., Einav et al., 2013; Handel, 2013; Bajari et al., 2014; Handel and Kolstad, 2015; Kowalski, 2015). Thus far, these models have focused almost exclusively on annual medical expenditure decisions, abstracting from several important features of an individual's optimization problem.

This paper contributes to the literature an estimate of insurance-induced moral hazard derived from a dynamic model of within-year medical care decisions under uncertainty. The model features an annual health insurance decision, followed by a sequence of medical care decisions made over the course of a health insurance year. The model allows medical care consumption to alter future health outcomes, assumes medical care prices are unknown prior to consumption, and allows for uncertainty about future medical care price, consumption, and health. I find that the distribution of insurance-induced spending produced by the model has thinner tails and a slightly larger mean than that produced a representative annual expenditure model. The key feature of the model that leads to differential effects is the presence of uncertainty at the time of a medical care decision. In a typical annual expenditure model, there is no uncertainty about the costs or benefits of medical care at the time that a medical care decision is made, which lessens the value of insurance, leading an individual to respond less strongly to insurance acquisition. The model is also used to (i) offer empirical support for the Grossman (1972) health capital model and (ii) explore quantitatively both the spending and welfare implications of a general shift towards HDHPs in the US.

Within a broader literature on structural estimation, this research highlights a potential consequence of aggregating decision variables (e.g., annual expenditure) in a way that necessarily prohibits the structural model from that accurately characterizing the incentives faced by and information available to an individual who, in reality, makes decisions with higher frequency.[47] In this research, the studied consequence is a biased predicted response to an environmental change in simulation, though other consequences may exist as well.

[47] A frequently modeled aggregate decision variable is annual savings (e.g., French and Jones, 2011). In reality, an individual makes multiple savings decisions over the course of the year while responding to changing interest rates, forecasting future changes in interest rates, and responding to expenditure shocks.

References

Akerlof, G. A., "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics* 84 (1970), 488–500.

Arcidiacono, P., H. Sieg and F. Sloan, "Living rationally under the volcano? An empirical analysis of heavy drinking and smoking," *International Economic Review* 48 (2007), 37–65.

Aron-Dine, A., L. Einav and A. Finkelstein, "The RAND Health Insurance Experiment, Three Decades Later," *Journal of Economic Perspectives* 27 (2013), 197–222.

Aron-Dine, A., L. Einav, A. Finkelstein and M. Cullen, "Moral Hazard in Health Insurance: Do Dynamic Incentives Matter?," *Review of Economics and Statistics* 97 (2015), 725–741.

Arrow, K. J., "Uncertainty and the Welfare Economics of Medical Care," *The American Economic Review* 53 (1963), 941–973.

Bajari, P., H. Hong, A. Khwaja and C. Marsh, "Moral Hazard, Adverse Selection, and Health Expenditures: A Semiparametric Analysis," *RAND Journal of Economics* 45 (2014), 747–763.

Bellman, R., *Dynamic Programming* (Princeton: Princeton University Press, 1957).

Blau, D. and D. Gilleskie, "The Role of Retiree Health Insurance in the Employment Behavior of Older Men," *International Economic Review* 49 (2008), 475–514.

Brot-Goldberg, Z. C., A. Chandra, B. R. Handel and J. T. Kolstad, "What does a deductible do? The impact of cost-sharing on health care prices, quantities, and spending dynamics," *The Quarterly Journal of Economics* 132 (2017), 1261–1318.

Cardon, J. H. and I. Hendel, "Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey," *RAND Journal of Economics* 32 (2001), 408–427.

Center for Disease Control and Prevention, "Health, United States, 2011: with Special Feature on Socioeconomic Status and Health," (May 2012), DHHS Publication No. 2012 - 1232.

Cohen, A. and L. Einav, "Estimating risk preferences from deductible choice," *The American Economic Review* 97 (2007), 745–788.

Coughlin, T. A., J. Holahan, K. Caswell and M. McGrath, "Uncompensated care for the uninsured in 2013: A detailed examination," *The Urban Institute* (2014).

Cronin, C. J., M. Forsstrom and N. Papageorge, "Mental Health, Human Capital, and the Labor Market," *Working Paper* (2017).

Cutler, D. and S. Reber, "Paying for Health Insurance: The Trade-Off between Competition and Adverse Selection," *The Quarterly Journal of Economics* 113 (1998), 433–466.

Cutler, D. and R. Zeckhauser, "The Anatomy of Health Insurance," *Handbook of Health Economics* 1 (2000), 563–643.

Dafny, L. and J. Gruber, "Public insurance and child hospitalizations: access and efficiency effects," *Journal of Public Economics* 89 (2005), 109–129.

Dalton, C. M., G. Gowrisankaran and R. Town, "Salience, Myopia, and Complex Dynamic Incentives: Evidence from Medicare Part D," *NBER Working Paper* 21104 (2017).

De Nardi, M., E. French and J. B. Jones, "Medicaid insurance in old age," *The American Economic Review* 106 (2016), 3480–3520.

Einav, L., A. Finkelstein, S. Ryan, P. Schrimpf and M. Cullen, "Selection on Moral Hazard in Health Insurance," *American Economic Review* 103 (2013), 178 – 219.

Einav, L., A. Finkelstein and P. Schrimpf, "The response of drug expenditure to nonlinear contract design: evidence from medicare part D," *The Quarterly Journal of Economics* 130 (2015), 841–899.

Ellis, R. P., "Rational Behavior in the Presence of Coverage Ceilings and Deductibles," *RAND Journal of Economics* 17 (1986), 158–175.

Finkelstein, A., S. Taubman, B. Wright, M. Bernstein, J. Gruber, J. Newhouse, H. Allen, K. Baicker, and the Oregon Health Study Group, "The Oregon Health Insurance Experiment: Evidence from the First Year," *The Quarterly Journal of Economics* 127 (2012).

Fox, J. T., K. I. Kim, S. P. Ryan and P. Bajari, "A simple estimator for the distribution of random coefficients," *Quantitative Economics* 2 (2011), 381–418.

French, E. and J. B. Jones, "The Effects of Health Insurance and Self-Insurance on Retirement Behavior," *Econometrica* 79 (2011), 693–732.

Gilleskie, D. B., "A Dynamic Stochastic Model of Medical Care Use and Work Absence," *Econometrica* 66 (1998), 1–45.

Grossman, M., "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy* 80 (1972), 223–255.

Guilkey, D. K. and P. M. Lance, "Program Impact Estimation with Binary Outcome Variables: Monte Carlo Results for Alternative Estimators and Empirical Examples," *Studies in Applied and Empirical Econometrics* A Festschrift in Honor of Peter Schmidt (New York: Springer) (2013).

HANDEL, B., "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts," *The American Economic Review* 103 (2013), 2643–2682.

HANDEL, B. R. AND J. T. KOLSTAD, "Health insurance for "humans": Information frictions, plan choice, and consumer welfare," *The American Economic Review* 105 (2015), 2449–2500.

HECKMAN, J. AND B. SINGER, "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica* 52 (1984), 271–320.

KEANE, M. P. AND K. I. WOLPIN, "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," *The Review of Economics and Statistics* 76 (1994), 648–672.

KEELER, E. B., J. P. NEWHOUSE AND C. E. PHELPS, "Deductibles and the Demand for Medical Care Services: The Theory of a Consumer Facing a Variable Price Schedule under Uncertainty," *Econometrica* 45 (1977), 641–656.

KEELER, E. B. AND J. E. ROLPH, "The Demand for Episodes of Treatment in the Health Insurance Experiment," *Journal of Health Economics* 7 (1988), 337 – 367.

KHWAJA, A., "Estimating Willingness to Pay for Medicare using a Dynamic Life-Cycle Model of Demand for Health Insurance," *Journal of Econometrics* 156 (2010), 130 – 147.

KOLSTAD, J. AND A. KOWALSKI, "The Impact of Health Care Reform on Hospital and Preventive Care: Evidence from Massachusetts," *Journal of Public Economics* 96 (2012), 909–929.

KOWALSKI, A., "Estimating the Tradeoff Between Risk Protection and Moral Hazard with a Nonlinear Budget Set Model of Health Insurance," *International Journal of Industrial Organization* 43 (2015), 122–135.

LIN, H. AND D. W. SACKS, "Intertemporal Substitution in Health Care Demand: Evidence from the RAND Health Insurance Experiment," *NBER Working Paper* 22802 (2017).

LO SASSO, A. T., L. A. HELMCHEN AND R. KAESTNER, "The Effects of Consumer-Directed Health Plans on Health Care Spending," *Journal of Risk and Insurance* 77 (2010), 85–103.

MANNING, W. G., H. L. BAILIT, B. BENJAMIN AND J. P. NEWHOUSE, "The demand for dental care: evidence from a randomized trial in health insurance," *The Journal of the American Dental Association* 110 (1985), 895–902.

MANNING, W. G., J. P. NEWHOUSE, N. DUAN, E. B. KEELER AND A. LEIBOWITZ, "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment," *The American Economic Review* 77 (1987), 251–277.

Mroz, T. A., "Discrete Factor Approximations in Simultaneous Equation Models: Estimating the Impact of a Dummy Endogenous Variable on a Continuous Outcome," *Journal of Econometrics* 92 (1999), 233 – 274.

Newhouse, J. P. and R. C. I. E. Group, *Free for all?: lessons from the RAND health insurance experiment* (Harvard University Press, 1993).

Painter, M. and M. Chernew, *Counting Change: Measuring Health Care Prices, Costs, and Spending* (Princeton: Robert Wood Johnson Foundation, 2012).

Pauly, M. V., "The Economics of Moral Hazard: Comment," *The American Economic Review* 58 (1968), 531–537.

Rosenthal, J. A., X. Lu and P. Cram, "Availability of consumer prices from US hospitals for a common surgical procedure," *JAMA Internal Medicine* 173 (2013), 427–432.

Rust, J., "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica* 55 (1987), pp. 999–1033.

Towers Watson, "High-Performance Insights: Best Practices in Health Care," *20th Annual Towers Watson/National Business Group on Health Employer Survey on Purchasing Value in Health Care* (2016).

Zweifel, P. and W. G. Manning, "Moral hazard and consumer incentives in health care," *Handbook of Health Economics* 1 (2000), 409–459.